

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 11:38:01 ; Search time 84 Seconds  
(without alignments)  
610.341 Million cell updates/sec

Title: US-09-992-430B-22  
Perfect score: 1625  
Sequence: 1 MFQDTKSQAVRTDAKTVKV.....ELMEKSASQIKSVIEHLEIN 323

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5 ,

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_19Jun03: \*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: \*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: \*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: \*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: \*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: \*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: \*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: \*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: \*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: \*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: \*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: \*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: \*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: \*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: \*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: \*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: \*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: \*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: \*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: \*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: \*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: \*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: \*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: \*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1614.5	99.4	324	AAE25510	Kluyveromyces ther
2	952.5	58.6	313	AAE25511	Torulaspora pretor
3	663	40.8	319	AAE23265	Bacillus strain LN
4	654	40.2	319	AAE10591	L-lactic acid dehy
5	635	39.1	322	AAG82159	S. epidermidis ope
6	635	39.1	322	AAG82645	S. epidermidis ope
7	635	39.1	330	ABP39483	Staphylococcus epi
8	627	38.6	314	ABB53683	Lactococcus lactis
9	622.5	38.3	314	AAY25997	B. flavum lactate

10	622.5	38.3	314	22	AAG92956	C glutamicum prote
11	622.5	38.3	314	22	AAB79307	Corynebacterium gl
12	617	38.0	313	23	ABB47671	Listeria monocytog
13	599	36.9	316	20	AAY20054	B. burgdorferi ant
14	596	36.7	327	23	ABP27743	Streptococcus poly
15	594.5	36.6	298	11	AAR08355	Recombinant thermo
16	594.5	36.6	298	11	AAR09295	Recombinant thermo
17	594	36.6	332	22	ABB64908	Drosophila melanog
18	589.5	36.3	325	23	ABB54670	Lactococcus lactis
19	576.5	35.5	328	12	AAR15736	L-lactic acid dehy
20	576.5	35.5	328	24	ABU01649	S. pneumoniae type
21	576.5	35.5	332	24	ABP81596	Streptococcus pneu
22	576	35.4	329	23	ABP27742	Streptococcus poly
23	573	35.3	332	9	AAP80891	Testis-specific la
24	568	35.0	353	21	AAG30183	Arabidopsis thalia
25	562	34.6	332	22	AAM40272	Human polypeptide
26	562	34.6	332	23	ABP65131	Hypoxia-regulated
27	562	34.6	372	22	AAM42058	Human polypeptide
28	550.5	33.9	381	22	AAG89135	Human secreted pro
29	550.5	33.9	381	23	ABG32424	Human secreted pro
30	550.5	33.9	381	23	AAU11432	Human testicular l
31	548.5	33.8	295	20	AAY20055	B. burgdorferi ant
32	546	33.6	334	24	ABR47513	Breast cancer asso
33	546	33.6	334	24	ABU57627	Differentially exp
34	546	33.6	334	24	ABU07467	Protein differenti
35	546	33.6	354	22	ABG20193	Novel human diagno
36	546	33.6	354	22	ABG21596	Novel human diagno
37	545	33.5	333	18	AAW33108	Chicken lactic aci
38	543	33.4	333	23	ABG32957	Chicken lactate de
39	543	33.4	333	23	ABG32958	Chicken lactate de
40	542	33.4	327	23	ABP66043	Bifidobacterium lo
41	542	33.4	333	18	AAW33107	Chicken lactic aci
42	542	33.4	333.	23	ABG32959	Chicken lactate de
43	542	33.4	338	22	AAU49896	Propionibacterium
44	535	32.9	333	10	AAP91948	Pig H4 isoenzyme.
45	526	32.4	301	21	AAG30184	Arabidopsis thalia

ALIGNMENTS

RESULT 1  
AAE25510  
ID AAE25510 standard; Protein; 324 AA.  
XX  
AC AAE25510;  
XX  
DT 30-OCT-2002 (first entry)  
XX Kluyveromyces thermotolerans lactate dehydrogenase.  
DE Lactate dehydrogenase; lactic acid; fermentation; organic product;  
KW food preservative; pharmaceutical product; cosmetic product; LDH;  
KW food additive; plastic product.  
XX  
OS Kluyveromyces thermotolerans.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 189  
FT /note= "This residue is absent in the sequence  
FT shown as SEQ ID NO:22 in sequence listing"  
XX  
PN WO200242471-A2.  
XX  
PD 30-MAY-2002.  
XX  
PF 23-NOV-2001; 2001WO-US44041.  
XX  
PR 22-NOV-2000; 2000US-252541P.  
XX  
PA (CRGI ) CARGILL DOW POLYMERS LLC.  
XX Rajgarhia V;

XX WPI; 2002-537387/57.  
DR N-PSDB; AAD41577.  
XX  
PT New nucleic acid molecule encoding yeast lactate dehydrogenase protein  
PT useful for transforming yeast cells for producing lactic acid -  
XX  
XX  
PS Claim 1; Page 69-71; 157pp; English.  
XX  
CC The invention relates to nucleic acid molecules encoding yeast lactate  
CC dehydrogenase (LDH) protein, especially Kluyveromyces thermotolerans or  
CC Torulaspora pretoriensis LDH protein. Polynucleotides of the invention  
CC are useful for transforming yeast cells for producing lactic acid by  
CC fermentation. They are used for producing organic products e.g. esters,  
CC carboxylates, carbohydrates, alditols, aldehydes, lactones, vitamins,  
CC alkenes, alkynes, lipids, amino acids which are useful as preservatives  
CC or additives in food, pharmaceutical or cosmetic products. They are  
CC also used to make plastic and other products. The present sequence is  
CC Kluyveromyces thermotolerans LDH protein.  
XX  
SQ Sequence 324 AA;  
Query Match 99.4%; Score 1614.5; DB 23; Length 324;  
Best Local Similarity 99.7%; Pred. No. 2.7e-142;  
Matches 323; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1 MFQDTKSOAVRTDAKTVKVVVGVSGSATAYTLLSGIVSEIVLIDVNDKKAEGESMD 60  
Db 1 MFQDTKSOAVRTDAKTVKVVVGVSGSATAYTLLSGIVSEIVLIDVNDKKAEGESMD 60  
QY 61 LNHAAPSNTSRAGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKY 120  
Db 61 LNHAAPSNTSRAGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKY 120  
QY 121 APDTILLIATNPVDVLTYSYKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDA 180  
Db 121 APDTILLIATNPVDVLTYSYKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDA 180  
QY 181 CVIGEHDG-GVPVWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIKRGY 239  
Db 181 CVIGEHDGSGVPVWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIKRGY 240  
QY 240 TSYGIAAGLLRIVKAILEDTGSTLTSTVGVGYFGVEQIAISVPTKLNKSGAHQVAELSLD 299  
Db 241 TSYGIAAGLLRIVKAILEDTGSTLTSTVGVGYFGVEQIAISVPTKLNKSGAHQVAELSLD 300  
QY 300 EXEIELMEKSASQIKSVIEHLEIN 323  
Db 301 EXEIELMEKSASQIKSVIEHLEIN 324  
RESULT 2  
AAE25511 standard; Protein; 313 AA.  
XX  
AC AAE25511;  
DT 30-OCT-2002 (first entry)  
XX  
DE Torulaspora pretoriensis lactate dehydrogenase.  
XX  
KW Lactate dehydrogenase; lactic acid; fermentation; organic product;  
KW food preservative; pharmaceutical product; cosmetic product; LDH;  
KW food additive; plastic product.  
XX  
OS Torulaspora pretoriensis.  
XX  
PN WO200242471-A2.  
XX  
PD 30-MAY-2002.  
XX  
PF 23-NOV-2001; 2001WO-US44041.  
XX

PR 22-NOV-2000; 2000US-252541P.  
XX (CRGI ) CARGILL DOW POLYMERS LLC.  
PA  
XX Rajgarhia V;  
XX  
DR WPI; 2002-537387/57.  
DR N-PSDB; AAD41584.  
XX  
PT New nucleic acid molecule encoding yeast lactate dehydrogenase protein  
PT useful for transforming yeast cells for producing lactic acid -  
XX  
XX Claim 2; Page 144-146; 157pp; English.  
XX  
CC The invention relates to nucleic acid molecules encoding yeast lactate  
CC dehydrogenase (LDH) protein, especially Kluyveromyces thermotolerans or  
CC Torulaspora pretoriensis LDH protein. Polynucleotides of the invention  
CC are useful for transforming yeast cells for producing lactic acid by  
CC fermentation. They are used for producing organic products e.g. esters,  
CC carboxylates, carbohydrates, alditols, aldehydes, lactones, vitamins,  
CC alkenes, alkynes, lipids, amino acids which are useful as preservatives  
CC or additives in food, pharmaceutical or cosmetic products. They are  
CC also used to make plastic and other products. The present sequence is  
CC Torulaspora pretoriensis LDH protein.  
XX  
SQ Sequence 313 AA;  
Query Match 58.6%; Score 952.5; DB 23; Length 313;  
Best Local Similarity 60.1%; Pred. No. 1.8e-80;  
Matches 187; Conservative 55; Mismatches 66; Indels 3; Gaps 2;  
QY 15 KTVKVVVGVSGVSATAYTLLSGIVSEIVLIDVNDKKAEGESMDLNHAAPSNTSR-- 72  
Db 3 RCAKVAIVGAGQVGSTTAYTLLSSLVAEVVLIDVDRKVEGQFMDLNHAAPLTKESTRS 62  
QY 73 AGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNP 132  
Db 63 AGDYESCADAADVIVTGGANQKPGQTRMELAEARNVKIMQEVIPKIVKYAPNAILLIATNP 122  
QY 133 VDVLTYISYKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG-GVP 191  
Db 123 VDVLTYASLKASGFPASRVIGSGTVLDSARIQHNLSKLFNVSSVNAFIIGEHGDSSVP 182  
QY 192 VWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIKRGYTSYGIAAGLLRI 251  
Db 183 VWSLAEIAGMKVEDYCRQSKRKFDPSILTKIYEESRDAAYIIERKGTNFGIAAGLARI 242  
QY 252 VKAILEDGTSTLTSTVGVGYFGVEQIAISVPTKLNKSGAHQVAELSDEKEIEMEKSAS 311  
Db 243 VRAILRDEGALLTVSTVGEHFGMKDVSLSVPTRVDRSGAHVVDLLNDKELEIKTSGA 302  
QY 312 QIKSVIEHLEI 322  
Db 303 KIKSACDELGI 313  
RESULT 3  
AAE23265  
ID AAE23265 standard; Protein; 319 AA.  
XX  
AC AAE23265;  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE Bacillus strain LN lactate dehydrogenase (LDH).  
XX  
KW Thermophilic; recombination; gene mutation; ethanol production;  
KW bacterial fermentation; lactate dehydrogenase; LDH; enzyme.  
XX  
OS Bacillus sp. TN.  
XX  
PN WO200229030-A2.  
XX

PD 11-APR-2002.  
XX  
PF 05-OCT-2001; 2001WO-GB04434.  
XX  
PR 06-OCT-2000; 2000GB-0024554.  
PR 13-NOV-2000; 2000US-247017P.  
XX  
PA (ELSW-) ELSWORTH BIOTECHNOLOGY LTD.  
XX  
PI Javed M, Cusdin F, Milner P, Green E;  
XX  
DR WPI; 2002-426110/45.  
DR N-PSDB; AAD37417.  
XX  
PT Novel thermophilic gram-positive bacterium which has been transformed  
PT by homologous recombination for stabilizing a gene mutation and for  
PT inserting an expressible gene, useful for ethanol production -  
XX  
PS Claim 26; Fig 6; 44pp; English.  
XX  
CC The invention relates to a thermophilic gram-positive bacterium which  
CC has been transformed using a method of homologous recombination for  
CC stabilising a gene mutation and for inserting an expressible gene.  
CC The thermophilic gram positive bacterium is useful for continuous  
CC production of ethanol by bacterial fermentation, using optimised  
CC fermentation conditions, where the pH of the fermentation medium is  
CC within the range of pH 5.5-7.5, preferably 6.4-6.9, the temperature of  
CC the fermentation medium is within the range 40-75 degrees, preferably  
CC 60-68 degrees, and the feed dilution rate is 0.3-0.8 per hour, preferably  
CC 0.4-0.6 per hour. The method involves using air sparging within the  
CC culture such that the redox potential is between -360 to -400 mV,  
CC preferably 370 to -380 mV. The present sequence is Bacillus strain LN  
XX lactate dehydrogenase (LDH).  
SQ Sequence 319 AA;  
Query Match 40.8%; Score 663; DB 23; Length 319;  
Best Local Similarity 44.6%; Pred. No. 2.1e-53;  
Matches 136; Conservative 59; Mismatches 104; Indels 6; Gaps 4;  
QY 18 KVVVGVGSGSATAYTLLSGIVSEIVLIDVNDKKAEGESMDLNHA---APSNTRSRAG 74  
Db 8 RVALIGTFVGASYAFALMNQGIADLVLDVNDKKAEGDVMDLNHGKVFAPKPMNIWFG 67  
QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
Db 68 DYQDCQADLVVICAGANQKPGETRLDLVDKNINIFKTIIVDSVMKSGFDGVLVATNPVD 127  
QY 135 VLTYISYKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEGHD-GVPVW 193  
Db 128 ILTYATWKFSGLPKERVIGSGTILDTARFRLLSEYFQVAPTNVHAYIIGEGHDTLPVW 187  
QY 194 SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYTSYGIAGLLRIVK 253  
Db 188 SHAEIGSIPVEQILMQ-NDNYRKEDLDNIFVNVVDAAYQIIKKGATYYGIAMGLVRITR 246  
QY 254 AILEDGTSTLTVST-VGDYFGVEQIAISVPTKLNKSGAHQVAELSDEKEIEMKESASQ 312  
Db 247 AILHNENAILTVSAHLDCGYGERNVYIGVPAIINRNGIREVMELTLNETEQQFHHSVTV 306  
QY 313 IKSVI 317  
Db 307 LKDIL 311  
RESULT 4  
AAR10591  
ID AAR10591 standard; Protein; 319 AA.  
XX  
AC AAR10591;  
XX  
DT 25-MAR-2003 (updated)  
DT 20-MAR-1991 (first entry)

XX L-lactic acid dehydrogenase.  
DE  
XX LDH; L-lactic acid; thermophile.  
KW  
XX Bacillus strain TP-262.  
OS  
XX JP02286077-A.  
PN  
XX 26-NOV-1990.  
PD  
XX 27-APR-1989; 89JP-0108432.  
PF  
XX 27-APR-1989; 89JP-0108432.  
PR  
XX (TOYA ) TOYAMA CHEM CO LTD.  
PA  
XX WPI; 1991-012213/02.  
DR N-PSDB; AAQ10162.  
DR  
XX Bacillus sp. pref. TP-262 stain - is thermophilic bacterium  
PT producing L-lactic acid dehydrogenase.  
PT  
XX Disclosure; Fig 1; 16pp; Japanese.  
PS  
XX The sequence was deduced from DNA which can be used to produce  
CC L-LDH by recombinant DNA techniques. The bacillus species contg.  
CC the DNA and able to express it is new and is thermophilic, growing  
CC at 50-70 deg.C.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
CC  
SQ Sequence 319 AA;  
Query Match 40.2%; Score 654; DB 12; Length 319;  
Best Local Similarity 43.9%; Pred. No. 1.5e-52;  
Matches 134; Conservative 59; Mismatches 106; Indels 6; Gaps 4;  
QY 18 KVVVGVGSGSATAYTLLSGIVSEIVLIDVNDKKAEGESMDLNHA---APSNTRSRAG 74  
Db 8 RVALIGTFVGASYAFALMNQGIADLVLDVNDKKAEGDVMDLNHGKVFAPKPMNIWHG 67  
QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
Db 68 DYQDCQADLVVICAGANQKPGETRLDLVNDKNMNIKTIIVDSVMKSGFDGIFLVATNPVD 127  
QY 135 VLTYISYKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEGHD-GVPVW 193  
Db 128 ILTYATWKFSGLPKERVIGSGTILDTARFRFLVSEYFQVAPTNVHAYIIGEGHDTLPVW 187  
QY 194 SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYTSYGIAGLLRIVK 253  
Db 188 SHAEIGSVPIEQILSQ-NDYRKEDLENI FNVNRDAAYQVIEKKGATYYGIAMGLVRITR 246  
QY 254 AILEDGTSTLTVSTVGD-YFGVEQIAISVPTKLNKSGAHQVAELSDEKEIEMKESASQ 312  
Db 247 AILHNENAILTVSAYLDGQYNEQNVYIGVPAIINRNGIREVMELKLNETEQQQFHHSVTV 306  
QY 313 IKSVI 317  
Db 307 LKDIL 311  
RESULT 5  
AAG82159  
ID AAG82159 standard; Protein; 322 AA.  
XX  
AC AAG82159;  
XX  
DT 03-SEP-2001 (first entry)  
XX  
DE S. epidermidis open reading frame protein sequence SEQ ID NO:1412.  
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;







QY	18	KVVVVGVGSVGSATAYTLLSGIVSEIVLIDVNDKKAEGESMDLNHA---	APSNTRSRAG	74
Db	7	KVVVIGTGFVGTSAIYSMINQGLVNLVLIDVNDKKAEGEALDLDGVSWGQENVIVRAG	66	
QY	75	DYDCCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD	134	
Db	67	DYKDCKNADIVVVTAGVNOQKPGQSRDLVNTNAKIMRSIVTQVMDSGFDGFVIAASN	PVD	126
QY	135	VLTYISYKASGFPLSRVIGSGTVLDTARPKVILGEHFKISSDSIDACVIGEHDG	-VPVW	193
Db	127	ILTYVAVWETSGLDQSRIVGTGTTDTRPRKELATKLEIDPRSVHGYIIGEHDSEVAVW	186	
QY	194	SLTNIDGMKLRDYCEKANHIFDQNAFHRIPEQTRDAAAYDIIKRKGYSYAGIAGLLRIVK	253	
Db	187	SHTTVGGKPILEFIVNKKIGVED-LSNLSNKVKNAAYEIIDKKQATYYIGIGMSTARIVK	245	
QY	254	AILEDTGSTLTVSTV--GDYFGVEQIAISVPTKLNKSGAHQVAELSDEKEIELMEKSAS	311	
Db	246	AILNNEQAILPVSAYLARGEY-GQEGVFTGVPFSIVNQNGVREIIELNIDAYEKKQFEKSVS	304	
QY	312	QIKSVIEHLE	321	
Db	305	QLKEVIESIK	314	

RESULT 9	
AAAY25997	
ID	AAAY25997 standard; Protein; 314 AA.
XX	
AC	AAAY25997;
XX	
DT	18-OCT-1999 (first entry)
XX	
DE	B. flavum lactate dehydrogenase protein.
XX	
KW	Lactate dehydrogenase; destruction; lactic acid.
XX	
OS	Brevibacterium flavum.
XX	
PN	JP11206385-A.
XX	
PD	03-AUG-1999.
XX	
PF	28-JAN-1998; 98JP-0030594.
XX	
PR	28-JAN-1998; 98JP-0030594.
XX	
PA	(MITU ) MITSUBISHI CHEM CORP.
XX	
DR	WPI; 1999-486360/41.
DR	N-PSDB; AAZ09139.
XX	
PT	New lactate dehydrogenase gene and strain destroying gene - useful
PT	for formation of lactic acid
XX	
PS	Claim 1; Page 11-12; 13pp; Japanese.

This invention describes the isolation of a novel lactate dehydrogenase protein from *Brevibacterium flavum* strain MJ-233. The invention also describes a microbial strain destroying the lactate dehydrogenase gene of a microbe in which the lactate dehydrogenase gene is destroyed by a homologous recombination of the above DNA or the above recombinant vector DNA with the lactate dehydrogenase gene on the chromosome DNA of the microbe cell and a method for the preparation of an amino acid or an organic acid in which the above strain destroying the lactate dehydrogenase gene is cultured in a medium and the amino acid or the organic acid (except lactic acid) is collected from the culture. Formation of lactic acid in the preparation of an amino acid and an organic acid can be decreased with no control of oxygen concentration during culture. This sequence represents the lactate dehydrogenase protein described in the specification.

Query Match	38.3%	Score 622.5	DB 20	Length 314
Best Local Similarity	43.9%	Pred. No. 1.3e-49		
Matches 134	Conservative 54	Mismatches 108	Indels 9	Gaps 5
18	KVVWVGSGSATAYTLLSGIVSEIVLIDVKNKDAEGESMDLNHA--APSNTRSAG	74		
Db	8	KIVLIGADVGVAAYALINQGMADHLAIIDIDEKKLEGVMDLNHGVMWADSRTRVTKG	67	
Qy	75	DYDPCAGAAIVIVTCGINKQNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD	134	
Db	68	TYADCEDAAMVVICAGAAQKPGETRLQLVDKNVKIMKSIVGDVNASGDFGIFLVASNPVD	127	
Qy	135	VLTYISYKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDISACVIGEHD-GVPVW	193	
Db	128	ILTYAVWKFSGLEWNRVIGSGTVLDSARFRYMLGELYEVAPSSVHAYIIEGHGDTLPVL	187	
Qy	194	SLTNIDGMKLRDYCEKANHFIDQNAFHRIFEQTRDAAYDIIKRKGYTSYGIAAGLLRIVK	253	
Db	188	SSATIAGVSLSRMLDKDPDL--EGRLEKIFEDTRDAAYHIIIDAKGSTSYGIGMLARITR	245	
Qy	254	AILEDTGSTLTVSTV--GDYFGVEQIAISVPTKLNKSGAHQVAELSDKEIEMKESAS	311	
Db	246	AILQNDVAVPVSALLHGEY-GEEDIYIGTPAVVNRGIRRVVELEITDHEMERFKHSAN	304	
Qy	312	QIKSV	316	
Db	305	TLREI	309	

RESULT 10  
AAG92956  
ID AAG92956 standard; Protein; 314 AA.  
XX  
AC AAG92956;  
XX  
DT 26-SEP-2001 (first entry)  
XX  
DE C glutamicum protein fragment SEQ ID NO: 6710.  
XX  
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
KW organic acid synthesis.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN EP1108790-A2.  
XX  
PD 20-JUN-2001.  
XX  
PF 18-DEC-2000; 2000EP-0127688.  
XX  
PR 16-DEC-1999; 99JP-0377484.  
PR 07-APR-2000; 2000JP-0159162.  
PR 03-AUG-2000; 2000JP-0280988.  
XX  
PA (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi  
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
XX  
DR WPI; 2001-376931/40..  
DR N-PSDB; AAH68175.  
XX  
PT Novel polynucleotides derived from Coryneform bacteria, for ident  
PT mutation point of a gene, measuring expression of a gene, analys  
PT expression profile or pattern of a gene and identifying homolog  
PT  
XX  
PS Claim 17; SEQ ID NO: 6710; 246pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of nucleotide and protei  
CC sequences from the Coryneform bacterium Corynebacterium glutamic

CC are useful for identifying the mutation point of a gene derived from a  
CC mutant of corynebacterium, measuring expression amount and  
CC analysing the expression profile or expression pattern of a gene derived  
CC from Corynebacterium, and identifying a homologue of a gene derived  
CC from corynebacterium. Corynebacterium bacteria are useful for producing  
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
CC particularly L-lysine. The present sequence is a protein described  
CC in the exemplification of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC European Patent Office.  
XX  
SQ Sequence 314 AA;

Query Match 38.3%; Score 622.5; DB 22; Length 314;  
Best Local Similarity 43.9%; Pred. No. 1.3e-49;  
Matches 134; Conservative 54; Mismatches 108; Indels 9; Gaps 5;

QY 18 KVVVGVGSGSATAYTLLLSGIVSEIVLIDVNVKDKAEGESMDLNHA---APSNTSRAG 74  
Db 8 KIVLIGAGDVGVAAYALINQGMADHLAIIDIDEKKLEGNVMDLNHGVVWADSRTRVTKG 67  
QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
Db 68 TYADCEDAAMVVICAGAAQKPGETRLQLVDKNVKIMKSIIVGDMDSGFGDIFLVASNPVD 127  
QY 135 VLTYSISKASGPELSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG-GVPVW 193  
Db 128 ILTYAVWKFSGLEWNRVIGSGTVLDSARFRYMLGELYEVAPSSVHAYIIGEGDTELPVL 187  
QY 194 SLTNIDGMKLRDYCEKANHFIDQNAFHRIFEOTRDAAYDIIKRKYTSYGIAAGLLRIVK 253  
Db 188 SSATIAGVSLSRMLDKDPEL--EGRLEKIFEDTRDAAYHIIIDAKGSTSYGIGMGLARITR 245  
QY 254 AILEDTGSTLTVSTV--GDYFGEVQIAISVPTKLNKSGAHQVAELSLDEKEIEMKXSAS 311  
Db 246 AILQNQDVAVPSALLHGEY-GEEDIYIGTFAVVRNRGIRRVVELEITDHEMERFKHSAN 304  
QY 312 QIKSV 316  
Db 305 TIREI 309

RESULT 11  
AAB79307  
ID AAB79307 standard; Protein; 314 AA.  
XX AAB79307;  
AC AAB79307;  
XX 30-APR-2001 (first entry)  
DT Corynebacterium glutamicum SMP protein sequence SEQ ID NO:130.  
XX Corynebacterium glutamicum; carbon metabolism and energy production;  
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;  
KW fine chemical production; organic acid; proteinogenic amino acid;  
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;  
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;  
KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;  
KW diagnosis; Corynebacterium diphtheriae; evolutionary study.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN WO200100844-A2.  
XX  
PD 04-JAN-2001.  
XX  
XX 23-JUN-2000; 2000WO-IB00943.  
PF  
XX  
PR 25-JUN-1999; 99US-0141031.  
PR 08-JUL-1999; 99DE-1031412.  
PR 08-JUL-1999; 99DE-1031413.  
PR 08-JUL-1999; 99DE-1031419.

PR 08-JUL-1999; 99DE-1031420.  
PR 08-JUL-1999; 99DE-1031424.  
PR 08-JUL-1999; 99DE-1031428.  
PR 08-JUL-1999; 99DE-1031431.  
PR 08-JUL-1999; 99DE-1031433.  
PR 08-JUL-1999; 99DE-1031434.  
PR 08-JUL-1999; 99DE-1031510.  
PR 08-JUL-1999; 99DE-1031562.  
PR 08-JUL-1999; 99DE-1031634.  
PR 09-JUL-1999; 99DE-1032180.  
PR 09-JUL-1999; 99DE-1032227.  
PR 09-JUL-1999; 99DE-1032230.  
PR 09-JUL-1999; 99US-0143208.  
PR 14-JUL-1999; 99DE-1032924.  
PR 14-JUL-1999; 99DE-1032973.  
PR 14-JUL-1999; 99DE-1033005.  
PR 27-AUG-1999; 99DE-1040765.  
PR 31-AUG-1999; 99US-0151572.  
PR 03-SEP-1999; 99DE-1042076.  
PR 03-SEP-1999; 99DE-1042079.  
PR 03-SEP-1999; 99DE-1042086.  
PR 03-SEP-1999; 99DE-1042087.  
PR 03-SEP-1999; 99DE-1042088.  
PR 03-SEP-1999; 99DE-1042095.  
PR 03-SEP-1999; 99DE-1042123.  
PR 03-SEP-1999; 99DE-1042125.  
XX  
XX (BADI ) BASF AG.

PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;  
XX  
XX WPI; 2001-061975/07.  
DR N-PSDB; AAF71424.  
XX

PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar  
PT metabolism and oxidative phosphorylation protein for production or  
PT modulation of production of fine chemicals e.g. amino acids,  
PT carbohydrates or enzymes -

PS Claim 20; Page 333-334; 1246pp; English.

XX  
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar  
CC metabolism and oxidative phosphorylation (SMP) proteins given in  
CC AAB79243 to AAB 79633 which are involved in carbon metabolism and  
CC energy production. The C. glutamicum SMP gene can be used in vectors  
CC (II) for expression in host cells and production or modulation of  
CC production of fine chemicals, such as, an organic acid, a proteinogenic  
CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,  
CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty  
CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a  
CC cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins  
CC (III) encoded by them are used for diagnosing the presence or activity of  
CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells  
CC containing them are used to map genomes of organisms related to  
CC C. glutamicum, identify and localise C. glutamicum sequences of interest,  
CC in evolutionary studies, in determining SMP protein regions required  
CC for function, in modulating SMP protein activity, in modulating the  
CC metabolism of sugars, and in modulating high-energy molecule production  
CC in a cell (i.e. ATP, NADPH).

XX Sequence 314 AA;

Query Match 38.3%; Score 622.5; DB 22; Length 314;  
Best Local Similarity 43.9%; Pred. No. 1.3e-49;  
Matches 134; Conservative 54; Mismatches 108; Indels 9; Gaps 5;

QY 18 KVVVGVGSGSATAYTLLLSGIVSEIVLIDVNVKDKAEGESMDLNHA---APSNTSRAG 74  
Db 8 KIVLIGAGDVGVAAYALINQGMADHLAIIDIDEKKLEGNVMDLNHGVVWADSRTRVTKG 67  
QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
Db 68 TYADCEDAAMVVICAGAAQKPGETRLQLVDKNVKIMKSIIVGDMDSGFGDIFLVASNPVD 127









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 11:42:17 ; Search time 29 Seconds  
(without alignments)  
471.255 Million cell updates/sec

Title: US-09-992-430B-22  
Perfect score: 1625  
Sequence: 1 MFQDTKSQAVRTDAKTVKV.....ELMEKSASQIKSVIEHLEIN 323

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	648	39.9	317	1	US-08-748-068-3
2	635	39.1	330	4	US-09-134-001C-4328
3	599	36.9	324	4	US-09-107-532A-6486
4	578	35.6	331	4	US-09-711-681-4
5	569.5	35.0	307	4	US-09-107-532A-4274
6	562.5	34.6	327	1	US-08-748-068-2
7	550.5	33.9	381	4	US-09-711-681-2
8	545	33.5	333	2	US-08-869-506-3
9	545	33.5	333	3	US-09-128-967-3
10	542	33.4	333	2	US-08-869-506-2
11	542	33.4	333	3	US-09-128-967-2
12	535	32.9	333	1	US-08-748-068-1
13	492	30.3	320	3	US-09-535-381-2
14	425	26.2	329	1	US-08-270-013B-2
15	425	26.2	329	1	US-08-838-418-2
16	395	24.3	325	4	US-09-134-001C-5533
17	383	23.6	315	5	PCT-US94-03796-2
18	329	20.2	330	3	US-08-676-882-2
19	324	19.9	304	4	US-09-634-238-238
20	280.5	17.3	148	4	US-09-370-838-215
21	179	11.0	338	3	US-08-922-957-3
22	178.5	11.0	338	3	US-08-922-957-1
23	164.5	10.1	298	3	US-08-922-957-4
24	151.5	9.3	95	4	US-09-724-623-108
25	143	8.8	273	4	US-09-091-097-12
26	134.5	8.3	327	1	US-08-211-682-25
27	103.5	6.4	453	1	US-08-374-155A-8

28	103.5	6.4	453	2	US-08-785-396-8	Sequence 8, Appli
29	103	6.3	3892	4	US-09-328-352-5503	Sequence 5503, Ap
30	98.5	6.1	906	4	US-09-328-352-6037	Sequence 6037, Ap
31	97.5	6.0	474	4	US-08-311-731A-353	Sequence 353, App
32	97.5	6.0	725	4	US-08-311-731A-105	Sequence 105, App
33	97	6.0	684	4	US-09-328-352-5781	Sequence 5781, Ap
34	96	5.9	302	4	US-09-107-532A-5280	Sequence 5280, Ap
35	95.5	5.9	251	4	US-09-648-004-20	Sequence 20, Appl
36	95	5.8	558	4	US-09-071-035-268	Sequence 268, App
37	95	5.8	1638	4	US-09-071-035-258	Sequence 258, App
38	95	5.8	1638	4	US-09-071-035-262	Sequence 262, App
39	95	5.8	1638	4	US-09-071-035-266	Sequence 266, App
40	94	5.8	509	4	US-09-499-302A-10	Sequence 10, Appl
41	93.5	5.8	422	3	US-08-961-083-116	Sequence 116, App
42	93.5	5.8	422	4	US-09-536-784-116	Sequence 116, App
43	93	5.7	893	4	US-09-328-352-6626	Sequence 6626, Ap
44	92	5.7	267	4	US-09-773-748-1	Sequence 1, Appli
45	91.5	5.6	444	4	US-08-940-572-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-748-068-3  
; Sequence 3, Application US/08748068  
; Patent No. 5770410  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Chiral Synthesis  
; NUMBER OF SEQUENCES: 15  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/748,068  
; FILING DATE: 12-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/256,959  
; FILING DATE: 05-OCT-1994  
; APPLICATION NUMBER: GB 92 02033.8  
; FILING DATE: 30-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 92 04702.6  
; FILING DATE: 04-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 93/00204  
; INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 317 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (15^16)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (28^29)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (29^30)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:

NAME/KEY: Modified-site  
LOCATION: (30^31)  
OTHER INFORMATION: /note= "---- numbering  
OTHER INFORMATION: discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (40^41)  
OTHER INFORMATION: /note= "-- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (69^70)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (73^74)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (83^84)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (115^116)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (138^139)  
OTHER INFORMATION: /note= "---- numbering  
OTHER INFORMATION: discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (176^177)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (181^182)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (196^197)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (211^212)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (230^231)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (234^235)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (251^252)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (264^265)  
OTHER INFORMATION: /note= "---- numbering  
OTHER INFORMATION: discontinuity"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (286^287)  
OTHER INFORMATION: /note= "- numbering discontinuity"  
US-08-748-068-3

Query Match 39.9%; Score 648; DB 1; Length 317;  
Best Local Similarity 44.9%; Pred. No. 4e-60;  
Matches 137; Conservative 51; Mismatches 111; Indels 6; Gaps 4;  
QY 18 KVVVVGSGSATAYTLLLSGIVSEIVLIDVNMKKAEGESMDLNHAAA---APSNTRSRAG 74

Db 8 RVVVGAGFVGASYVFALMNQIADIEIVLIDANESKAIGDAMDFNHGKVFAPKPVDIWHG 67  
QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAANKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
Db 68 DYDDCRDADLVVICAGANQKPGETRLDLVDKNIAIFRSIVESYNASGFGQLFLVATNPVD 127  
QY 135 VLTYSYKASGFPPLSRVIGSGTGLDTRARFKYILGEHFKISSDSIDACVIGEGD-GVPVW 193  
Db 128 ILTYATWKFSGLPHERVIGSGTGLDTRARFKYILGEHFKISSDSIDACVIGEGD-GVPVW 187  
QY 194 SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIKRGYTSYGIAGLLRIVK 253  
Db 188 SQAYIGVMPIRKLVEKGE-EAQKDLERIFVNVDRDAAYQIIEKKGATYYGIAMGLARVTR 246  
QY 254 AILEDTCSTLTSTVGD-YFGEQIAISVPTKLNKSGAHQVAELSLDEKEIEMKESASQ 312  
Db 247 AILHNENAILTVSAYLDGLYGERDVYIGVPAVINRNGIREVIEIELNDDEKXRFHHSAA 306  
QY 313 IKSVI 317  
Db 307 LKSVL 311  
RESULT 2  
US-09-134-001C-4328  
; Sequence 4328, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4328  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4328

Query Match 39.1%; Score 635; DB 4; Length 330;  
Best Local Similarity 43.1%; Pred. No. 1e-58;  
Matches 137; Conservative 57; Mismatches 100; Indels 24; Gaps 6;  
QY 18 KVVVVGSGSATAYTLLLSGIVSEIVLIDVNMKKAEGESMDLNHAAA---PSNTRSRAG 74  
Db 21 KVLVGDGSGSSYAFAMVTQGADEFVIIDIAKDKVEADVVDLNLHGALYSSSPVTVKAG 80  
QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAANKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
Db 81 EYEDCKDADLVVITAGAPQKPGETRLQVLEKTKIMKSIIVTSVMSDGFDFFLIAANPVD 140  
QY 135 VLTYSYKASGFPPLSRVIGSGTGLDTRARFKYILGEHFKISSDSIDACVIGEGD-GVPVW 193  
Db 141 ILTRYVKEVTGLPAERVIGSGTGLDTRARFKYILGEHFKISSDSIDACVIGEGD-GVPVW 200  
QY 194 SLTNIDGMKLRDYCE-----KANHIFDQNAFHRIFEQTRDAAYDIKRGYTSYGIAG 246  
Db 201 SQANVGISVVDTLKEETGSDAKAN-----EYINTRDAAYDIIOAKGSTYYGIAL 251  
QY 247 GLLRIVKAILLEDTGSTLTSTVST-VGDYFGEQIAISVPTKLNKSGAHQVAELSLDEKEI 305  
Db 252 ALLRISKALLNNENSILTVSSQLNGQYGFNDVYLGTLPTLINQNGAVKIYETPLNDNELQL 311  
QY 306 MEKSASQIKSV---IEHL 320  
Db 312 LEKSVKLTEDTYDSIKHL 329



RESULT 3  
US-09-107-532A-6486  
; Sequence 6486, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 6486:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 324 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...324  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6486:  
US-09-107-532A-6486

Query Match 36.9%; Score 599; DB 4; Length 324;  
Best Local Similarity 42.5%; Pred. No. 6.3e-55;  
Matches 130; Conservative 61; Mismatches 107; Indels 8; Gaps 5;  
QY 18 KVVVGVGSGSATAYTLLSGIVSEIVLIDVKNKKAEGESMDLNHA--APSNTSRAG 74  
Db 17 KVVIVGTGFVGTSTAYAMINQGISNELVLIDVQKAEAGEALDLDGMAWGDNVAVWSG 76  
QY 75 DYPDCAGAAIVIVTCGINKQNGQTRMDLAANKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
Db 77 GYECKDADIVVITAGINQKPGQSRDLVKTNASIMRQIVKEIMGSGFDGIIVVASNPVD 136  
QY 135 VLTYSISKASGFPISRIVSGTVDLTARFKYILGEHFKISSDSIDACVIGEHDG-VPVW 193  
Db 137 ILTYIAWNESGLPTSRVIGTGTTLDTTRFRKEIALKLKVDPRSVHGVILGEHGDSEVAW 196  
QY 194 SLTNIDGMKLRDYCEKANHFIDQNAFHRIFEQTRDAAAYDIKRGYTSYGIAGLLRIVK 253  
Db 197 SHTTVGGKPVFEIVK-DHRIAQDELVDIADKVRNAAYEIIDRKKATYYIGMSTARIVK 255

QY 254 AILEDGTSTLTVST--VGDFYFGEQIAISVPTKLNKSGAHQVAELSLDEKEIELMEKSAS 311  
Db 256 AILNNEQAVLPVSAYLTGEY-DEKDIFTGVPISVDENGVRVVELSINEEEKAMFSKSTS 314  
QY 312 QIKSVI 317  
Db 315 ALREVL 320  
RESULT 4  
US-09-711-681-4  
; Sequence 4, Application US/09711681  
; Patent No. 6503743  
; GENERAL INFORMATION:  
; APPLICANT: LADUNGA, Steven et al.  
; TITLE OF INVENTION: Isolated Human Secreted Proteins,  
; Nucleic Acid Molecules Encoding Human Secreted Proteins And  
; TITLE OF INVENTION: Uses Thereof  
; FILE REFERENCE: CL000839  
; CURRENT APPLICATION NUMBER: US/09/711,681  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-711-681-4  
Query Match 35.6%; Score 578; DB 4; Length 331;  
Best Local Similarity 38.5%; Pred. No. 1.1e-52;  
Matches 120; Conservative 73; Mismatches 109; Indels 10; Gaps 5;  
QY 18 KVVVGVGSGSATAYTLLSGIVSEIVLIDVKNKKAEGESMDLNHAA----PSNTRSR 72  
Db 22 KITVVGAVGAVMACAISILMKDLADELALVDVMDKLGEMMDLQHGSLFLKTPKIVSSK 81  
QY 73 AGDYPDCAGAAIVIVTCGINKQNGQTRMDLAANKNANIMLEIIPNVAKYAPDTILLIATNP 132  
Db 82 --DYCVTANSKLVIITAGARQEGESRLNVQRNVNIFKFIIPNIVKYSPhCKLLIVSNP 139  
QY 133 VDVLTVISYKASGFPISRIVSGTVDLTARFKYILGEHFKISSDSIDACVIGEHDG-GVP 191  
Db 140 VDILTVYVANKISGFPKRVIGSGCNLDLSARFRLMGERLGVHLSCHGWVGEHGDSSVP 199  
QY 192 VWSLTNIDGMKLRDYCEKANHFIDQNAFHRIFEQTRDAAAYDIKRGYTSYGIAGLLRI 251  
Db 200 VWSGVNVAGVSLKSLNPELGTADKEQWKEVHKQVDSAYEVIKLGKGYTSAIGLSVADL 259  
QY 252 VKAILEDGTSTLTVST-VGDFYFV-EQIAISVPTKLNKSGAHQVAELSLDEKEIELMEKS 309  
Db 260 AESIMKNLRRVHPISITMIKGLYGINEDVFLSVPCILGQNGISDVVKVTLTPPEEARLKS 319  
QY 310 ASQIKSVIEHLE 321  
Db 320 ADTLWGIQKELQ 331  
RESULT 5  
US-09-107-532A-4274  
; Sequence 4274, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA



```

; LOCATION: (188^189)
; OTHER INFORMATION: /note= "- numbering discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (193^194)
; OTHER INFORMATION: /note= "- numbering discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (208^209)
; OTHER INFORMATION: /note= "- numbering discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (243^244)
; OTHER INFORMATION: /note= "- numbering discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (247^248)
; OTHER INFORMATION: /note= "- numbering discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (259^260)
; OTHER INFORMATION: /note= "- numbering discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (262^263)
; OTHER INFORMATION: /note= "- numbering discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (275^276)
; OTHER INFORMATION: /note= "---- numbering
; OTHER INFORMATION: discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (298^299)
; OTHER INFORMATION: /note= "- numbering discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (308^309)
; OTHER INFORMATION: /note= "- numbering discontinuity"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (326^327)
; OTHER INFORMATION: /note= "- numbering discontinuity"
US-08-748-068-2
```

```

Query Match      34.6%; Score 562.5; DB 1; Length 327;
Best Local Similarity 39.4%; Pred. No. 4.7e-51;
Matches 125; Conservative 67; Mismatches 116; Indels 9; Gaps 7;

QY 10 VRTDAKTVKVVVGVSGVSATAYTLLSGIVSEIVLIDVNDKKAEGESMDLNHAAP--S 67
Db 12 IEDDESQCKITIVGTGAVGMACAISILLKDLADELALVDVALDKLGEMMDLQHGSLFFS 71

QY 68 NTRSRAG-DYDPCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTIL 126
Db 72 TSKVTSGKDISVSANSRIVIVTAGARQGEQETRLALVQRNVAIMKIIPAIHVHSPDCKI 131

QY 127 LIATNPVDVLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEH 186
Db 132 LVVSNPVDILTIVWKISGLPVTTRVIGSGCNLDSARFYLIGELGVHPTSCHGWIIGEH 191

QY 187 GD-GVPVWSLTNIDGMKLRDYCEKANHFIDQNAFHRIFEQTRDAAYDIKRGYTSYGIA 245
Db 192 GDSSVPLWSGVNVAGVALKTLDPKLGTDSDKEHWKNIHQVIOQSAIEIILKGYTSSWAIG 251

QY 246 AGLLRIVKAILEDTGSTLTVST-VGDYFGV-EQIAISVPTKLNKSGAHQVAELSDEKEI 303
Db 252 LSVMDLVP--LKNLRRVHPVSTMVKGLYGIKEBELFLSIPCVLGRNGVSDVVKIDLSEEE- 308

QY 304 ELMKESASQIKSVIEHL 320
Db 309 ALLKKAETLWNIQKNL 325
```

```

RESULT 7
US-09-711-681-2
; Sequence 2, Application US/09711681
; Patent No. 6503743
; GENERAL INFORMATION:
; APPLICANT: LADUNGA, Steven et al.
; TITLE OF INVENTION: Isolated Human Secreted Proteins,
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Secreted Proteins And
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: CL000839
; CURRENT APPLICATION NUMBER: US/09/711,681
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-711-681-2
```

```

Query Match      33.9%; Score 550.5; DB 4; Length 381;
Best Local Similarity 36.3%; Pred. No. 1.1e-49;
Matches 119; Conservative 76; Mismatches 122; Indels 11; Gaps 5;

QY 6 KSQAVR--TDAKTV---KVVVVGVSGVSATAYTLLSGIVSEIVLIDVNDKKAEGESMD 60
Db 54 KSELIERFTSEKPVHHSKVSIIIGTGSVGMACAISILLKGLSDELALVDLDEKLGKGETMD 113

QY 61 LNHAAPSNTSR--AGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNV 117
Db 114 LQHGSPFTKMPNIVCSKDYFVTANSNLVITAGARQEKETRLNLVQRNVAFKLMISSI 173

QY 118 AKYAPDTILLIATNPVDVLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDS 177
Db 174 VOYSPHCKLIIVSNPVDILTIVAWKLSAFFPKNRIIGSGCNLDTARFRFLIGKLGHSSES 233

QY 178 IDACVIGEHGD-GVPVWSLTNIDGMKLRDYCEKANHFIDQNAFHRIFEQTRDAAYDIK 236
Db 234 CHGWILGEHGDSSVPVWSGVNIAGVPLKDLNSDIGTDKDPQEWKNVHKEVTATAYEIIKM 293

QY 237 KGYTSYGIAAGLLRIVKAILEDTGSTLTVSTV--GDYFGVEQIAISVPTKLNKSGAHQVA 294
Db 294 KGYTSSWAIGLSVADLTESILKNLRIHPVSTITKGLYGIDEVEFLSIPCILGENGITNLI 353

QY 295 ELSLDEKEIELMEKSASQIKSVIEHLEI 322
Db 354 KIKLTPEEEAHLKKSATKLWEIQNKLL 381
```

```

RESULT 8
US-08-869-506-3
; Sequence 3, Application US/08869506
; Patent No. 5827710
; GENERAL INFORMATION:
; APPLICANT: Uchida, Kohji
; APPLICANT: Matsukawa, Hirokazu
; APPLICANT: Matuo, Yushi
; APPLICANT: Fujita, Tutosi
; TITLE OF INVENTION: DNA ENCODING THE SUBUNIT OF AVIAN
; TITLE OF INVENTION: LACTATE DEHYDROGENASE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 No. 5827710th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```





; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201-4741  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/869,506  
; FILING DATE: 05-JUN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 73797/1996  
; FILING DATE: 29-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford, Arthur R.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 159-43  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 333 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-869-506-2

Query Match 33.4%; Score 542; DB 2; Length 333;  
Best Local Similarity 36.8%; Pred. No. 7.1e-49;  
Matches 114; Conservative 71; Mismatches 119; Indels 6; Gaps 4;

QY 18 KVVVGVGSGVFATYLLSGIVSEIVLIDVNDKKAEGESMDLNHAA---PSNTRSAG 74  
Db 22 KITVGVGVGMACAISILGKGLCDELALVDVLEDKLGEMMDLQHGSLFLQTHKIVADK 81  
QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
Db 82 DYAVTANSKIVVVTAGVRQOEGESRLNVQRNVNVFKFIIPQIVKYSFNCTILVVSNPVD 141  
QY 135 VLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG-GVPVW 193  
Db 142 ILTYVTWKLSQLPKHRVIGSGCNLDTARFRYLMAERLGIHPTSCHGWILGEHGDSSAVW 201  
QY 194 SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRGYTSYGIAAGLLRIVK 253  
Db 202 SGNVVGVSLQELNPFAMGTDKDSENWKEVHKQVVEAYEVIRLKGYNWAIGLSVAELCE 261  
QY 254 AILEDGTSTLTVST-VGDYFGVE-QIAISVPTKLNKSGAHQVAELSDEKEIEMEKSAS 311  
Db 262 TMLKNLYRVHSVSTLVKGTGYIENDVFLSLPCVLSASGLTSVINQKLKDDVAQLKKSAD 321  
QY 312 QIKSVIEHLE 321  
Db 322 TLWSIQDLK 331

RESULT 11  
US-09-128-967-2  
; Sequence 2, Application US/09128967  
; Patent No. 6057141  
; GENERAL INFORMATION:  
; APPLICANT: Uchida, Kohji  
; APPLICANT: Matsukawa, Hirokazu  
; APPLICANT: Matuo, Yushi  
; APPLICANT: Fujita, Tutosi  
; TITLE OF INVENTION: DNA ENCODING THE SUBUNIT OF AVIAN  
; TITLE OF INVENTION: LACTATE DEHYDROGENASE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 No. 6057141th Glebe Rd. 8th floor  
; CITY: Arlington  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201-4741  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/128,967  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/869,506  
; FILING DATE: 05-JUN-1997  
; APPLICATION NUMBER: JP 73797/1996  
; FILING DATE: 29-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford, Arthur R.  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 159-43  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4000  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 333 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-128-967-2

Query Match 33.4%; Score 542; DB 3; Length 333;  
Best Local Similarity 36.8%; Pred. No. 7.1e-49;  
Matches 114; Conservative 71; Mismatches 119; Indels 6; Gaps 4;

QY 18 KVVVGVGSGVSATAYTLLSGIVSEIVLIDVNDKKAEGESMDLNHAA---PSNTRSAG 74  
Db 22 KITVGVGVGMACAISILGKGLCDELALVDVLEDKLGEMMDLQHGSLFLQTHKIVADK 81  
QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
Db 82 DYAVTANSKIVVVTAGVRQOEGESRLNVQRNVNVFKFIIPQIVKYSFNCTILVVSNPVD 141  
QY 135 VLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG-GVPVW 193  
Db 142 ILTYVTWKLSQLPKHRVIGSGCNLDTARFRYLMAERLGIHPTSCHGWILGEHGDSSAVW 201  
QY 194 SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRGYTSYGIAAGLLRIVK 253  
Db 202 SGNVVGVSLQELNPFAMGTDKDSENWKEVHKQVVEAYEVIRLKGYNWAIGLSVAELCE 261  
QY 254 AILEDGTSTLTVST-VGDYFGVE-QIAISVPTKLNKSGAHQVAELSDEKEIEMEKSAS 311  
Db 262 TMLKNLYRVHSVSTLVKGTGYIENDVFLSLPCVLSASGLTSVINQKLKDDVAQLKKSAD 321  
QY 312 QIKSVIEHLE 321  
Db 322 TLWSIQDLK 331

RESULT 12  
US-08-748-068-1  
; Sequence 1, Application US/08748068  
; Patent No. 5770410  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Chiral Synthesis  
; NUMBER OF SEQUENCES: 15  
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/748,068  
;; FILING DATE: 12-NOV-1996  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/256,959  
;; FILING DATE: 05-OCT-1994  
;; APPLICATION NUMBER: GB 92 02033.8  
;; FILING DATE: 30-JAN-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: GB 92 04702.6  
;; FILING DATE: 04-MAR-1992  
;; APPLICATION DATA:  
;; APPLICATION NUMBER: GB 93/00204  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 333 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (29^30)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (42^43)  
;; OTHER INFORMATION: /note= "--- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (43^44)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (44^45)  
;; OTHER INFORMATION: /note= "---- numbering  
;; OTHER INFORMATION: discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (54^55)  
;; OTHER INFORMATION: /note= "-- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (83^84)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (87^88)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (97^98)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (129^130)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (152^153)  
;; OTHER INFORMATION: /note= "---- numbering  
;; OTHER INFORMATION: discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (190^191)

;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (195^196)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (210^211)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (245^246)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (249^250)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (266^267)  
;; OTHER INFORMATION: /note= "--- numbering discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (279^280)  
;; OTHER INFORMATION: /note= "---- numbering  
;; OTHER INFORMATION: discontinuity"  
;;  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: (302^303)  
;; OTHER INFORMATION: /note= "- numbering discontinuity"  
;;  
;; US-08-748-068-1  
;;  
Query Match 32.9%; Score 535; DB 1; Length 333;  
Best Local Similarity 35.3%; Pred. No. 3.9e-48;  
Matches 110; Conservative 74; Mismatches 118; Indels 10; Gaps 5;  
QY 18 KVVVGVGSGSATAYTLTLLSGIVSEIVLIDVNDKKAEGESMDLNHAA-----PSNTRSR 72  
Db 22 KITVVGQVGVGMACAISILGKSLTDLALVDVLEDKLKGEMMDLQHGSLFLQTPKIVANK 81  
QY 73 AGDYPDCAGAAIVIVTCGINQKNQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNP 132  
Db 82 --DYSVTANSKIVVVTAGVRQGEESRLNVQRNVNVFKFIIPQIVKYSNCIIIVVSNP 139  
QY 133 VDLTVIYKASGFPLSRVIGSGTVLDTARFKYILGEHFHKISSDSIDACVIGEGD-GVP 191  
Db 140 VDILTVYTWKLSGLPKHRVIGSGCNLDSARFRYLMAEKLGVHPSSCHGWILGEHGDSSVA 199  
QY 192 VWSLTNIDGMKLRDYCEKANHFIDQNAFHRIFEQTRDAAYDIKRGYTSYGIAGLLRI 251  
Db 200 VWSGVNVAGVSLQQLNPEMGTDNDSENWKEVHKVGVESAYEVIKLGTYTNWAGLSVADL 259  
QY 252 VKAILEDGTSTLTVST-VGDYFVGE-QIAISVPTKLNKSGAHQVAELSLDEKEIELMEKS 309  
Db 260 IESMLKNLSRIHPVSTMVQGMVIGIENEVFLSLPCVLNARGLTSVINQKLKDEVAQLKNS 319  
QY 310 ASQIKSVIEHLE 321  
Db 320 ADTLWGIQKDLK 331  
RESULT 13  
US-09-535-381-2  
; Sequence 2, Application US/09535381  
; Patent No. 6268189  
; GENERAL INFORMATION:  
; APPLICANT: Skory, Christopher D  
; TITLE OF INVENTION: Fungal Lactate Dehydrogenase Gene and Constructs for  
; TITLE OF INVENTION: the Expression Thereof  
; FILE REFERENCE: Dkt 0110.98 - Christopher D. Skory  
; CURRENT APPLICATION NUMBER: US/09/535,381  
; CURRENT FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 8



; SEQUENCE CHARACTERISTICS:

;  
; LENGTH: 329 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-838-418-2

Query Match 26.2%; Score 425; DB 1; Length 329;

Best Local Similarity 35.8%; Pred. No. 1.7e-36;

Matches	107;	Conservative	58;	Mismatches	114;	Indels	20;	Gaps	9;
---------	------	--------------	-----	------------	------	--------	-----	------	----

QY 15 KTVKVVGVGVSATAYTLLSGIVSEIVLIDVN--KDKAEGESMDLNHAAP-----S 67

Db 4 KRKKISVIGAGFTGATTAP - LLAQKELGDVVLDIPQLENPTKKGKALDMEASPVLGFDA 62

QY 68 NTRSRAGDYPCAGAIVIVTCGINQKNQOTRMDLAAKNANIMLEIIPNVAKYAPDTILL 127

Db 63 NIIG-TSDYADTADSDIVVITAGIARKPGMSRDDLVTNNQIMKQVTKVEVVKYSPNCYII 121

QY 128 IATNPVDLTYISYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHG 187

27 122 VLTNPVDAMTYTVFKESGFPKRVICQSGVLDTARFRTFAAELNISKVDVTGTVLGGHG 181

QV 188 DG-VPVWSLTNIDGMKLRDYCEKANHIFDONAFHRIFEQTRDAAAYDIIKRKGYS--YGI 244

**Q7**      D G V F V S T I N I D G C H E R I C E M A L K I L D V A T K R I L I N G S I Y P    100  
         ||| : | : |     ||| : :: ||| : : :     |     |     |  
**Q8**      D G V F V S T I N I D G C H E R I C E M A L K I L D V A T K R I L I N G S I Y P    100  
         ||| : | : |     ||| : :: ||| : : :     |     |     |  
**b6**      D P M V P L V R Y S Y A G G I P L -----E K-----L I P K O R L D A I V E R T R K G G E I V N L L N G S A Y Y A P    234

OV 245 AAGLLRIVKAILEDTGSTL-TVSTVGDYFGVEOIAISVPTKLNKSGAHOVAELSLDEKE 302

[illegible]

Search completed: July 25, 2003, 11:47:21

```

Job time : 30 secs

```





db	128	ILTYATWKFSGLPKERVIGSGTILDTARFRFLLSYFQVAPTNNVHAYII GEHGDTLPVM	187
QY	194	SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYSYGI AAGLLRIVK	253
Db	188	SHABIGSIPVEQIILMQ -NDNYRKEDLDNIFVNVRDAAAYQII EKKGATYYGIAMGLVRITR	246
QY	254	AILEDTGSTLTVST -VG DYFGVEQIAISVPTKLNKSGAHQVAELSDEKEI ELMKSASQ	312
Db	247	AILHNENAILTVSAHLDGQYGERNVYIGVPAIINRNGIREVMELTTLNETEQQFFHHSVTV	306
QY	313	IKSVI	317
Db	307	LKDIL	311

## RESULT 2

US-09-971-361-9  
; Sequence 9, Application US/09971361  
; Patent No. US20020081677A1  
; GENERAL INFORMATION:  
; APPLICANT: Javed, Muhammad  
; APPLICANT: Cusdin, Fiona  
; APPLICANT: Milner, Paul  
; APPLICANT: Green, Edward  
; TITLE OF INVENTION: Ethanol Production  
; FILE REFERENCE: 000487.00010  
; CURRENT APPLICATION NUMBER: US/09/971,361  
; CURRENT FILING DATE: 2001-12-28  
; PRIOR APPLICATION NUMBER: US 60/247,017  
; PRIOR FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: UK 0024554.8  
; PRIOR FILING DATE: 2000-10-06  
; NUMBER OF SEQ ID NOS: 16

ORGANISM: BACILLUS BULGARI IN  
US-09-971-361-9

Qy 313 IKSVI 317  
:|::  
Db 307 LKDIL 311

### RESULT 3

US-09-738-626-6710  
Sequence 6710, Application US/09738626

```

; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6710
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
;
US-09-738-626-6710

```

## RESULT 4

```

US-10-274-266-4
; Sequence 4, Application US/10274266
; Publication No. US20030059893A1
; GENERAL INFORMATION:
; APPLICANT: LADUNGA, Steven et al.
; TITLE OF INVENTION: Isolated Human
; TITLE OF INVENTION: Secreted Prot
; FILE REFERENCE: CL000839DIV
; CURRENT APPLICATION NUMBER: US/10/
; CURRENT FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Vers
.; SEQ ID NO 4

```

```

; LENGTH: 331
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-274-266-4

Query Match      35.6%; Score 578; DB 15; Length 331;
Best Local Similarity 38.5%; Pred. No. 4.9e-51;
Matches 120; Conservative 73; Mismatches 109; Indels 10; Gaps 5;

QY 18 KVVVVGSGSATAYTLLSGIVSEIVLIDVNDKKAEGESMDLNHAA-----PSNTRSR 72
Db 22 KITVVGGAAGMACAISILMKDLADELALVDVWEDKLGEMMDLQGSFLFKTPKIVSSK 81

QY 73 AGDYPCAGAAIVIVTCGINQKNGQTRMDLAAKNVANIMLEIIPNVAKYAPDTILLIATNP 132
Db 82 --DYCVTANSKLVIITAGARQQEGESRLNLVQRNVNIFKFIIPNIVKYSPhCKLLIVSNP 139

QY 133 VDVLTYISYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG-GVP 191
Db 140 VDILTYVAVKISGFPPKNRVIGSGCNLDSARFRYLMGERLGVHALSCHGWVLGEHGDSSVP 199

QY 192 VWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKYTSYGIAAGLLRI 251
Db 200 VMSGVNVAGSLKSLNPELGTADKEQWKEVHKQVWDSAYEVIKLKYTSWAIGLSVADL 259

QY 252 VKAILEDTGSTLTVST-VGDYFGV-EQIAISVPTKLNKSGAHQVAELSLDEKEIELMEKS 309
Db 260 AESIMKNLRRVHPISITMIKGLYGINEDVFLSVPCILQNGISDVVKVTLTPEEARLKKS 319

QY 310 ASQIKSVIEHLE 321
Db 320 ADTLWGIQKELQ 331
```

```

RESULT 5
US-09-731-872-255
; Sequence 255, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 255
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -33...-1
US-09-731-872-255
```

```

Query Match      33.9%; Score 550.5; DB 10; Length 381;
Best Local Similarity 36.3%; Pred. No. 4.1e-48;
Matches 119; Conservative 76; Mismatches 122; Indels 11; Gaps 5;

QY 6 KSQAVR--TDAKTV---KVVVVGSGSATATYLLSGIVSEIVLIDVNDKKAEGESMD 60
Db 54 KSELIERFTSEKPVHHSKVSIIIGTSVGMACAISILLKGLSDELALVDLDEDKLGKGETMD 113

QY 61 LNHAAPSNTSR---AGDYPCAGAAIVIVTCGINQKNGQTRMDLAAKNVANIMLEIIPNV 117
Db 114 LQHGSPFTKMPNIVCSKDYFVTANSNLVIITAGARQEKGETRLNLVQRNVAIFKLMISSI 173
```

```

QY 118 AKYAPDTILLIATNPVDVLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDS 177
Db 174 VQYSPHCKLIIIVSNPVDILTYVAVKLSAFPKNRIIGSGCNLDTARFRFLIGQKLGHS 233

QY 178 IDACVIGEHDG-GVPVWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKR 236
Db 234 CHGWILGEHGDSSVPVMSGVNIAGVPLKDLNSDIDGTDKDPEQWKNVHKEVTATAYEIIKM 293

QY 237 KGYTSYGIAAGLLRIVKAILEDTGSTLTIVSTV--GDYFGVEQIAISVPTKLNKSGAHQVA 294
Db 294 KGYTSWAIGLSVADLTESILKNLRIHPVSTIIKGLYGIDEEVFLSIPCILGENGITNLI 353

QY 295 ELSLDEKEIELMEKSASQIKSVIEHLEI 322
Db 354 KIKLTPEEEAHLKKSAKTLWEIQNKLL 381

RESULT 6
US-10-274-266-2
; Sequence 2, Application US/10274266
; Publication No. US20030059893A1
; GENERAL INFORMATION:
; APPLICANT: LADUNGA, Steven et al.
; TITLE OF INVENTION: Isolated Human Secreted Proteins, Nucleic Acid Molecules Encoding
; TITLE OF INVENTION: Secreted Proteins, And Uses Thereof
; FILE REFERENCE: CL000839DIV
; CURRENT APPLICATION NUMBER: US/10/274,266
; CURRENT FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-274-266-2
```

```

Query Match      33.9%; Score 550.5; DB 15; Length 381;
Best Local Similarity 36.3%; Pred. No. 4.1e-48;
Matches 119; Conservative 76; Mismatches 122; Indels 11; Gaps 5;

QY 6 KSQAVR--TDAKTV---KVVVVGSGSATATYLLSGIVSEIVLIDVNDKKAEGESMD 60
Db 54 KSELIERFTSEKPVHHSKVSIIIGTSVGMACAISILLKGLSDELALVDLDEDKLGKGETMD 113

QY 61 LNHAAPSNTSR---AGDYPCAGAAIVIVTCGINQKNGQTRMDLAAKNVANIMLEIIPNV 117
Db 114 LQHGSPFTKMPNIVCSKDYFVTANSNLVIITAGARQEKGETRLNLVQRNVAIFKLMISSI 173

QY 118 AKYAPDTILLIATNPVDVLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDS 177
Db 174 VQYSPHCKLIIIVSNPVDILTYVAVKLSAFPKNRIIGSGCNLDTARFRFLIGQKLGHS 233

QY 178 IDACVIGEHDG-GVPVWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKR 236
Db 234 CHGWILGEHGDSSVPVMSGVNIAGVPLKDLNSDIDGTDKDPEQWKNVHKEVTATAYEIIKM 293

QY 237 KGYTSYGIAAGLLRIVKAILEDTGSTLTIVSTV--GDYFGVEQIAISVPTKLNKSGAHQVA 294
Db 294 KGYTSWAIGLSVADLTESILKNLRIHPVSTIIKGLYGIDEEVFLSIPCILGENGITNLI 353

QY 295 ELSLDEKEIELMEKSASQIKSVIEHLEI 322
Db 354 KIKLTPEEEAHLKKSAKTLWEIQNKLL 381
```

```

RESULT 7
US-09-974-298-75
; Sequence 75, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
```

; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 75
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 3365683CD1
US-09-974-298-75

Query Match 33.6%; Score 546; DB 10; Length 334;
Best Local Similarity 36.2%; Pred. No. 9.7e-48;
Matches 113; Conservative 72; Mismatches 117; Indels 10; Gaps 5;

QY 18 KVVVVGVSATAYTLLSGIVSEIVLIDVNDKKAEGESMDLNHAA-----PSNTRSR 72
Db 23 KITVVGQVGMACAISILGKSLADELALVDLEDKLGEMMDLQHGSFLQTPKIVADK 82

QY 73 AGDYPDCAGAAIVITCGINQKNGOTRMDLAANKNANIMLEIIPNVAKYAPDTILLIATNP 132
Db 83 --DYSVTANSKIVVVTAGVRQEGESRLNLVQRNVNVFKFIIPQIVKYSPPDCIIIVVSNP 140

QY 133 VDVLTYISYKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHGD-GVP 191
Db 141 VDILTYVTWKLSGLPKHRVIGSGCNLDSARFRYLMAEKLGIHPSSCHGWILGEHGDSSVA 200

QY 192 VWSLTNIDGMKLRDYCEKANHFIDQNAFHRIFEQTRDAAYDIIKRKGYSYGYIAAGLLRI 251
Db 201 VWSGVNVAGVLSQELNPEIGTDNDSENKWEVHKMVVESAYEVIKLGKYNWAIGLSVADL 260

QY 252 VKAILEDTGSTLTVST-VGDYFGVE-QIAISVPTKLNKSGAHQVAELSDEKEIELMEKS 309
Db 261 IESMLKNLSRIHPVSTWVKMGYGIENEVFLSLPCILNARGLTSVINQKLKDDDEVAQLKKS 320

QY 310 ASQIKSVIEHLE 321
Db 321 ADTLWDIQKOLK 332

RESULT 8
US-10-177-293-262.
; Sequence 262, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887

; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 262
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-262

Query Match 33.6%; Score 546; DB 15; Length 334;
Best Local Similarity 36.2%; Pred. No. 9.7e-48;
Matches 113; Conservative 72; Mismatches 117; Indels 10; Gaps 5;

QY 18 KVVVVGVSATAYTLLSGIVSEIVLIDVNDKKAEGESMDLNHAA-----PSNTRSR 72
Db 23 KITVVGQVGMACAISILGKSLADELALVDLEDKLGEMMDLQHGSFLQTPKIVADK 82

QY 73 AGDYPDCAGAAIVITCGINQKNGOTRMDLAANKNANIMLEIIPNVAKYAPDTILLIATNP 132
Db 83 --DYSVTANSKIVVVTAGVRQEGESRLNLVQRNVNVFKFIIPQIVKYSPPDCIIIVVSNP 140

QY 133 VDVLTYISYKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHGD-GVP 191
Db 141 VDILTYVTWKLSGLPKHRVIGSGCNLDSARFRYLMAEKLGIHPSSCHGWILGEHGDSSVA 200

QY 192 VWSLTNIDGMKLRDYCEKANHFIDQNAFHRIFEQTRDAAYDIIKRKGYSYGYIAAGLLRI 251
Db 201 VWSGVNVAGVLSQELNPEMGTDNDSENKWEVHKMVVESAYEVIKLGKYNWAIGLSVADL 260

QY 252 VKAILEDTGSTLTVST-VGDYFGVE-QIAISVPTKLNKSGAHQVAELSDEKEIELMEKS 309
Db 261 IESMLKNLSRIHPVSTWVKMGYGIENEVFLSLPCILNARGLTSVINQKLKDDDEVAQLKKS 320

QY 310 ASQIKSVIEHLE 321
Db 321 ADTLWDIQKDLK 332

RESULT 9
US-09-971-536-51
; Sequence 51, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Bloksberg, Leonard
; APPLICANT: Lubbers, Mark
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul
; APPLICANT: Reid, Julian
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Methods
; TITLE OF INVENTION: Using Them
; FILE REFERENCE: 1043c2
; CURRENT APPLICATION NUMBER: US/09/971,536
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: PCT/NZ01/00160





US-10-144-649A-215

Query Match 17.3%; Score 280.5; DB 15; Length 148;  
Best Local Similarity 41.4%; Pred. No. 6e-21;  
Matches 53; Conservative 29; Mismatches 39; Indels 7; Gaps 2;

Qy 18 KVVVGVGVSATAYTLLSGIVSEIVLIDVNVKDKAEGESMDLNHAA-----PSNTRSR 72  
Db 23 KITVVGVGQVGMACAISILGKSLADELALVDVLEDKLKGEMMDLQHGSLFLQTPKIVADK 82

Qy 73 AGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNP 132  
Db 83 --DYSVTANSKIIVVVTAGVRQEGESRLNLVQRNVNVFKFIIPQIVKYSPODCIIIVVSNP 140

Qy 133 VDLVLYIS 140  
Db 141 VDILTYVT 148

RESULT 13

US-10-156-761-7888

; Sequence 7888, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 7888  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-7888

Query Match 13.8%; Score 223.5; DB 15; Length 303;  
Best Local Similarity 27.9%; Pred. No. 1.3e-14;  
Matches 75; Conservative 44; Mismatches 117; Indels 33; Gaps 9;

Qy 19 VVVVGVGVSATAYTLLSGIVSEIVLIDVNVKDKAEGESMDLNH-----AAPSNTRS-R 72  
Db 4 VGVVGAGAVGQTVAATLVASGICPRLLVSVRTVEQARALAADLDDMRQTGSPVQPEAR 63

Qy 73 AGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNP 132  
Db 64 VADLIGHAVVAVRAAFTNTRAADVVRMGGALTNPVIRALATTLRGY--QGTVLVWTNP 121

Qy 133 VDLVLYISYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGHGDGVPV 192  
Db 122 VDLMTRLFAETSGCP--RVYIGSNLDSARYRLTLAHLDDVPATTVGHVIGHGDGAVV 179

Qy 193 W-SLTNIDG---MKLRDYCEKANHFIDQNAFHRIFEQTRDAAYDIIKRKGYSYGIAG 247  
Db 180 CASSTTVNGTPAAVPLAE-----VRAELRTRPGQISAGVGRTRSGPAGA 223

Qy 248 LLRIIV-KAI-LEDTGSTLTVSTVGDYFGV 274  
Db 224 VLSTLRKALGLVDGTEELTAEHRGDWLGI 252

; Sequence 6816, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 6816  
; LENGTH: 146  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (90)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC FEATURE  
; LOCATION: (105)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC FEATURE  
; LOCATION: (114)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC FEATURE  
; LOCATION: (145)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-6816

Query Match 12.5%; Score 202.5; DB 15; Length 146;  
Best Local Similarity 38.9%; Pred. No. 6.3e-13;  
Matches 42; Conservative 22; Mismatches 37; Indels 7; Gaps 2;

Qy 18 KVVVGVGVSATAYTLLSGIVSEIVLIDVNVKDKAEGESMDLNHAA-----PSNTRSR 72  
Db 39 KITVVGVGQVGMACAISILGKSLADELALVDVLEDKLKGEMMDLHHGSLFLXTPKIVADK 98

Qy 73 AGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKY 120  
Db 99 --DYSVTAXSKIIVVVTAXVRQEGESRLNLVQRNVNVFKFIIPQIVKY 144

RESULT 15

US-09-925-301-1533

; Sequence 1533, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1533  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-1533

Query Match 11.5%; Score 186.5; DB 9; Length 142;  
Best Local Similarity 39.2%; Pred. No. 2.7e-11;  
Matches 40; Conservative 20; Mismatches 35; Indels 7; Gaps 2;

















Db 307 LKSVL 311

RESULT 10

A84142

L-lactate dehydrogenase lctE [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C;Accession: A84142

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: A84142

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-310 <STO>

A;Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07656.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

C;Superfamily: L-lactate dehydrogenase

Query Match 39.9%; Score 649; DB 2; Length 310;

Best Local Similarity 45.2%; Pred. No. 7.4e-42;

Matches 138; Conservative 53; Mismatches 100; Indels 14; Gaps 6;

QY 18 KVVVGVGSGSATAYTLLSGIVSEIVLIDVNVKDKAEGESMDLNHA--APSNTRSRAG 74

Db 8 KIVLVGTGAVGSSYAYALMNQGISDELILVDLNEEKAKGDVLDLNHSIVYAPSPMEIKFG 67

QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD 134

Db 68 SYEDCKDAALVVICAGAAQKPGETRLDLVHKNVGIFESIVGNIMKSGFNGIFLVATNPVD 127

QY 135 VLTYSISKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG-VPVW 193

Db 128 ILAYATWKFSGLPKERVIGSGTVLDTARFYLIGEEMNAAPTSMVHGVIIGEHDGSQLPVW 187

QY 194 SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAAYDIIRKKGVTYSYGIAAGLLRIVK 253

Db 188 SSATIAGTPI-----APRLTDEKK-QEIAENVRDAAKYKIIIEAKGATYYGIATGLARITR 240

QY 254 AILEDTGSTLTVSTV--GDYFGEQIAISVPTKLNKSGAHQVAELSLDEKEIELMEKSAS 311

Db 241 AILKNENVVLPVGTLLGEN-GHDDVYIGVPAINREGVRQVVELSLNDEEKEKFARSVE 299

QY 312 QIKSV 316

Db 300 TLKDI 304

RESULT 11

DEBSLF

L-lactate dehydrogenase (EC 1.1.1.27) [validated] - Bacillus stearothermophilus

C;Species: Bacillus stearothermophilus

C;Date: 03-Aug-1984 #sequence\_revision 07-Feb-1997 #text\_change 03-Nov-2000

C;Accession: A26053; A29704; A91726; A00353

R;Barstow, D.A.; Clarke, A.R.; Chia, W.N.; Wigley, D.; Sharman, A.F.; Holbrook, J.J.; Ab

Gene 46, 47-55, 1986

A;Title: Cloning, expression and complete nucleotide sequence of the Bacillus stearother

A;Reference number: A26053; MUID:87106842; PMID:3026926

A;Accession: A26053

A;Molecule type: DNA

A;Residues: 1-317 <BAR>

A;Cross-references: GB:M14788; NID:g143139; PIDN:AAA22568.1; PID:g143140

R;Zuelli, F.; Weber, H.; Zuber, H.

Biol. Chem. Hoppe-Seyler 368, 1167-1177, 1987

A;Title: Structure and function of L-lactate dehydrogenases from thermophilic and mesoph

stearothermophilus, Bacillus caldolyticus and Bacillus caldotenax.

A;Reference number: A29704; MUID:88050101; PMID:3675869

A;Accession: A29704

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-317 <ZUE>

A;Cross-references: GB:M19396; GB:M19386; NID:g143137; PIDN:AAA22567.1; PID:g143138

R;Wirz, B.; Suter, F.; Zuber, H.

Hoppe-Seyler's Z. Physiol. Chem. 364, 893-909, 1983

A;Title: Structure and function of L-lactate dehydrogenases from thermophilic and mesoph;

A;Reference number: A91726; MUID:84006440; PMID:6352452

A;Accession: A91726

A;Molecule type: protein

A;Residues: 1-107,'SE',110-112,'H',114-317 <WIR>

A;Experimental source: NCIB 8924

R;Tratschin, J.D.; Wirz, B.; Frank, G.; Zuber, H.

Hoppe-Seyler's Z. Physiol. Chem. 364, 879-892, 1983

A;Title: Structure and function of L-lactate dehydrogenases from thermophilic and mesoph

. Cyanogen bromide fragments and partial sequence.

A;Reference number: A91725; MUID:84006439; PMID:6618448

A;Contents: annotation; CNBr fragments; partial sequence

A;Note: strain NCIB 8924

C;Function:

A;Description: catalyzes the reversible oxidation of (S)-lactate to pyruvate by NAD+

C;Superfamily: L-lactate dehydrogenase

C;Keywords: NAD; oxidoreductase; tetramer

F:8-38/Region: beta-alpha-beta NAD nucleotide-binding fold

F:152,179/Active site: Asp, His #status predicted

Query Match 39.9%; Score 648; DB 1; Length 317;

Best Local Similarity 44.9%; Pred. No. 9.1e-42;

Matches 137; Conservative 51; Mismatches 111; Indels 6; Gaps 4;

QY 18 KVVVGVGSGSATAYTLLSGIVSEIVLIDVNVKDKAEGESMDLNHA--APSNTRSRAG 74

Db 8 RVVVIGAGFVGASYVFALMNQGIADLVIDANESKAIGDAMDFNHGKVFAKPVDIWHG 67

QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD 134

Db 68 DYDDCRDADLVVICAGANQKPGETRLDLVDXNIAIFRSIVESVNASGFQGLFLVATNPVD 127

QY 135 VLTYSIKASGFPPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG-GVPVW 193

Db 128 ILTYATWKFSGLPHERVIGSGTILDTARFRLLGEYFVSAPQNVHAYIIGEHDGTELPVW 187

QY 194 SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAAYDIIRKKGVTYSYGIAAGLLRIVK 253

Db 188 SQAYIGVMPIRKLVESKGE-EAQKDLEIFVNVRDAAAYQIIIEKKGATYYGIAMGLARVTR 246

QY 254 AILEDTGSTLTVSTVGD-YFGEQIAISVPTKLNKSGAHQVAELSLDEKEIELMEKSASQ 312

Db 247 AILHNENAILTVSAYLDGLYGERDVYIGVPAVINRNGIREVIEIELNDDEKNRFHHSAT 306

QY 313 IKSVI 317

Db 307 LKSVL 311

RESULT 12

D89787

L-lactate dehydrogenase [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C;Accession: D89787

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguci

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: D89787

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-317 <KUR>

A;Cross-references: GB:BA000018; PID:g13700156; PIDN:BAB41455.1; GSPDB:GN00149

A;Experimental source: strain N315









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 11:38:26 ; Search time 23. Seconds  
(without alignments)  
660.419 Million cell updates/sec

Title: US-09-992-430B-22  
Perfect score: 1625  
Sequence: 1 MFQDTKSQAVRTDAKTVKV.....ELMEKSASQIKSVIEHLEIN 323

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	838	51.6	330	1 LDH_SCHPO	Q9p7p7 schizosacch
2	733	45.1	313	1 LDH_CLOAB	Q97md1 clostridium
3	723.5	44.5	318	1 LDH_BACME	P00345 bacillus me
4	722	44.4	317	1 LDH_CLOPE	Q8xp62 clostridium
5	687.5	42.3	319	1 LDH_THEMEA	P16115 thermotoga
6	679	41.8	318	1 LDHP_BACPS	P14561 bacillus ps
7	673	41.4	319	1 LDHX_BACPS	P20619 bacillus ps
8	667	41.0	317	1 LDH_BACCL	P10655 bacillus ca
9	659	40.6	317	1 LDH_BACCL	Q59244 bacillus ca
10	649	39.9	310	1 LDH_BACHD	Q9k5z8 bacillus ha
11	648	39.9	317	1 LDH_BACST	P00344 bacillus st
12	647.5	39.8	317	1 LDH1_STAAM	Q99wy2 staphylococ
13	639.5	39.4	325	1 LDH_BACSU	P00343 lactobacill
14	639	39.3	320	1 LDH_BACSU	P13714 bacillus su
15	627	38.6	314	1 LDH2_LACLA	Q9ci14 lactococcus
16	622.5	38.3	314	1 LDH_CORGL	Q8nlno corynebacte
17	618.5	38.1	325	1 LDH_LACSK	P50934 lactobacill
18	617	38.0	313	1 LDH_LISIN	Q92f65 listeria in
19	617	38.0	313	1 LDH_LISMO	P33380 listeria mo
20	613	37.7	310	1 LDH_THEAQ	P13715 thermus aqu
21	606	37.3	316	1 LDH_BORBU	O51114 borrelia bu
22	604	37.2	310	1 LDH_THECA	P06150 thermus cal
23	598	36.8	319	1 LDH2_STAAM	Q8num9 staphylococ
24	596	36.7	319	1 LDH2_STAAM	Q99r35 staphylococ
25	596	36.7	326	1 LDH_STRPY	Q99zn5 streptococc
26	595	36.6	323	1 LDH_LACHE	Q32765 lactobacill
27	594	36.6	332	1 LDH_DROME	Q95028 drosophila
28	591.5	36.4	324	1 LDH_LACLC	P04034 lactococcus
29	589.5	36.3	324	1 LDH1_LACLA	Q01462 lactococcus
30	586.5	36.1	330	1 LDHA_DISEL	Q9pw61 dissostichu
31	586.5	36.1	330	1 LDHA_PARCR	Q93538 parachaenic
32	585.5	36.0	330	1 LDHA_DISMA	O93544 dissostichu
33	583	35.9	332	1 LDHA_SQUAC	P00341 squalus aca

RESULT 1  
LDH\_SCHPO  
ID LDH\_SCHPO STANDARD; PRT; 330 AA.  
AC Q9P7P7;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable L-lactate dehydrogenase (EC 1.1.1.27) (L-LDH).  
GN SPAC186.08C.

OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Warren T., Whitehead S.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).

CC -!- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.

CC -!- PATHWAY: Anaerobic glycolysis; final step.

CC -!- SUBUNIT: Homotetramer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: BELONGS TO THE LDH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

34 582.5 35.8 330 1 LDHA\_CHAAC 093619 chaenoceph  
35 582.5 35.8 330 1 LDHA\_CHAGU 093541 champsoceph  
36 582.5 35.8 330 1 LDHA\_LEPNU 093545 lepidonotot  
37 581.5 35.8 330 1 LDHA\_NOTAN 09prh8 notothenia  
38 581.5 35.8 330 1 LDHA\_PATTE 093546 patagonotot  
39 580.5 35.7 330 1 LDHA\_GOBGI 093540 gobionototh  
40 580.5 35.7 330 1 LDHA\_NOTCO 093539 notothenia  
41 580 35.7 333 1 LDHA\_AWEME 09w7m6 ambystoma m  
42 579 35.6 331 1 LDHA\_ELEMC 093542 eleginops m  
43 578.5 35.6 330 1 LDHA\_PARMG 093543 paranotothe  
44 578.5 35.6 331 1 LDHA\_CORNI 093401 coryphopter  
45 578 35.6 331 1 LDHA\_MOUSE P06151 mus musculu

ALIGNMENTS



RT "Structure and function of L-lactate dehydrogenases from thermophilic  
RT and mesophilic bacteria. VII. Nucleotide sequence of the lactate  
RT dehydrogenase gene from the mesophilic bacterium Bacillus megaterium.  
RT Preparation and properties of a hybrid lactate dehydrogenase  
RT comprising moieties of the B. megaterium and B. stearothermophilus  
RT enzymes.";  
RL Biol. Chem. Hoppe-Seyler 368:1391-1399(1987).  
RN [2]  
RP SEQUENCE.  
RX MEDLINE=88050100; PubMed=3118900;  
RA Stangl D., Wiederkehr F., Suter F., Zuber H.;  
RT "Structure and function of L-lactate dehydrogenases from thermophilic  
RT and mesophilic bacteria, V. The complete amino-acid sequence of the  
RT mesophilic L-lactate dehydrogenase from Bacillus megaterium.";  
RL Biol. Chem. Hoppe-Seyler 368:1157-1166(1987).  
RN [3]  
RP PRELIMINARY SEQUENCE.  
RA Wiederkehr F.;  
RL Thesis (1982), ETH Zurich, Switzerland.  
CC -!- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.  
CC -!- PATHWAY: Anaerobic glycolysis; final step.  
CC -!- SUBUNIT: Homotetramer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: BELONGS TO THE LDH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; M22305; AAA22566.1; -.  
DR PIR; S00133; DEBSLM.  
DR HSSP; P00344; ILDB.  
DR HAMAP; MF\_00488; -; 1.  
DR InterPro; IPR001557; L\_LDH.  
DR InterPro; IPR001236; ldh.  
DR InterPro; IPR000205; NAD\_binding.  
DR Pfam; PF00056; ldh; 1.  
DR Pfam; PF02866; ldh C; 1.  
DR PRINTS; PR00086; LLDHGRGNASE.  
DR PROSITE; PS00064; L\_LDH; 1.  
KW Oxidoreductase; NAD; Glycolysis.  
FT ACT\_SITE 181 181 BY SIMILARITY.  
SQ SEQUENCE 318 AA; 35035 MW; 17122190576E1485 CRC64;  
  
Query Match 44.5%; Score 723.5; DB 1; Length 318;  
Best Local Similarity 49.2%; Pred. No. 3.1e-46;  
Matches 150; Conservative 54; Mismatches 94; Indels 7; Gaps 5;  
  
QY 15 KTVKVVVGVGSGSATAYTLTLLSGIVSEIVLIDVNDKKAEGESMDLNHAAPSNTRSR-- 72  
DB 8 KTRKVAVIGTGFVGSYAFSMVNOQIANELVLIDMNKKAEGEARDINHGMPFATPMKIW 67  
  
QY 73 AGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDPTILLIATNP 132  
DB 68 AGDYKDCADADLAVITAGANQAPGETRLDLVEKNVKIFECIVKDIMNSGFDGIILVATNP 127  
  
QY 133 VDVLTYISYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEGD-GVP 191  
DB 128 VDILAHVTQKVGSLPNRVRIGSGTILDTARFRYLLSDYFEVDSRNVHAYIMGEHGDTEFP 187  
  
QY 192 VWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYTSYGIAAGLLRI 251  
DB 188 VWSHAQIGGVKLEHFINTA-AIEKEPDMQHLFEQTRDAAYHIIINRKGATYYGIAMGLVRI 246  
  
QY 252 VKAILEDTGSLTSTV--GDYFGVEQIAISVPTKLNKSGAHQVAELSLDEKIELMEKS 309  
DB 247 TKAILDDENSILTVSALLEGOY-GISDVYIGVPAINKNGVRQIIELNLTPEHQQLHS 305  
  
QY 310 ASQIK 314

Db 306 ASILK 310  
|| :|  
RESULT 4  
LDH\_CLOPE  
ID\_LDH\_CLOPE STANDARD; PRT; 317 AA.  
AC Q8XP62;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE L-lactate dehydrogenase (EC 1.1.1.27) (L-LDH).  
GN LDH OR CPE0103.  
OS Clostridium perfringens.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1502;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=13 / Type A;  
RX MEDLINE=21664373; PubMed=11792842;  
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,  
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;  
RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
RT flesh-eater.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).  
CC -!- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.  
CC -!- PATHWAY: Anaerobic glycolysis; final step.  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE LDH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AP003185; BAB79809.1; -.  
DR HAMAP; MF\_00488; -; 1.  
DR InterPro; IPR001557; L\_LDH.  
DR InterPro; IPR001236; ldh.  
DR InterPro; IPR000205; NAD\_binding.  
DR Pfam; PF00056; ldh; 1.  
DR Pfam; PF02866; ldh C; 1.  
DR PRINTS; PR00086; LLDHGRGNASE.  
DR PROSITE; PS00064; L\_LDH; 1.  
KW Oxidoreductase; NAD; Glycolysis; Complete proteome.  
FT ACT\_SITE 178 178 BY SIMILARITY.  
SQ SEQUENCE 317 AA; 34522 MW; 88C862C0B2B3D951 CRC64;  
  
Query Match 44.4%; Score 722; DB 1; Length 317;  
Best Local Similarity 45.2%; Pred. No. 4e-46;  
Matches 142; Conservative 69; Mismatches 97; Indels 6; Gaps 4;  
  
QY 15 KTVKVVVGVGSGSATAYTLTLLSGIVSEIVLIDVNDKKAEGESMDLNHAAP--SNTRSR 72  
DB 5 KTNKISIGAGFVGSTTAFALMQDGLASEIVIVDINKDKAHAEMDLAQGAAFVKSVDIK 64  
  
QY 73 AGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDPTILLIATNP 132  
DB 65 SGDYADTKDSDIVIIITAGVGPKEGTRLDIINKNLKIFQSIQVPEVVVYSPNSILLVVSNP 124  
  
QY 133 VDVLTYISYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEGD-GVP 191  
DB 125 VDILTYITYKLSGFPKERVIGSGTVLDTSRKYLMLSEHFDIDARNVHTYIIEGHGDSIT 184  
  
QY 192 VWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYTSYGIAAGLLRI 251  
DB 185 AWSLTNIAGANVEEYCKTVCANCDSGFKKELPEKVKNAAEYIINSKGYTNYAVAVATRI 244















```
QY      313 IKSVI 317
      :|||:
Db      307 LKSVL 311

RESULT 12
LDH1_STAAM      STANDARD;      PRT;      317 AA.
ID      LDH1_STAAM
AC      Q99WY2;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      L-lactate dehydrogenase 1 (EC 1.1.1.27) (L-LDH 1).
GN      LDHA OR LCTE OR SAV0241 OR SA0232 OR MW0217.
OS      Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS      Staphylococcus aureus (strain N315), and
OS      Staphylococcus aureus (strain MW2).
OC      Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX      NCBI_TaxID=158878, 158879, 196620;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MU50 / ATCC 700699, and N315;
RX      MEDLINE=21311952; PubMed=11418146;
RA      Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA      Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA      Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA      Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA      Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA      Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA      Hattori M., Ogasawara N., Hayashi H., Furumatsu K.;
RT      "Whole genome sequencing of methicillin-resistant Staphylococcus
RT      aureus.";
RN      Lancet 357:1225-1240(2001).
ID      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MW2;
RX      MEDLINE=22040717; PubMed=12044378;
RA      Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA      Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA      Yamamoto K., Hiramatsu K.;
RT      "Genome and virulence determinants of high virulence community-
RT      acquired MRSA.";
RN      Lancet 359:1819-1827(2002).
CC      -!- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.
CC      -!- PATHWAY: Anaerobic glycolysis; final step.
CC      -!- SUBUNIT: Homotetramer (By similarity).
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -!- SIMILARITY: BELONGS TO THE LDH FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AP003358; BAB56403.1; -.
DR      EMBL; AP003129; BAB41455.1; -.
DR      EMBL; AP004822; BAB94082.1; -.
DR      PIR; D89787; D89787.
DR      HSSP; P00344; 2LDB.
DR      HAMAP; MF_00488; -.
DR      InterPro; IPR001557; L_LDH.
DR      InterPro; IPR001236; ldh.
DR      InterPro; IPR000205; NAD binding.
DR      InterPro; IPR000594; ThiF domain.
DR      Pfam; PF00056; ldh; 1.
DR      Pfam; PF02866; ldh_C; 1.
DR      PRINTS; PR00086; LLDHDEGNASE.
DR      PROSITE; PS00064; L_LDH; 1.
KW      Oxidoreductase; NAD; Glycolysis; Complete proteome.
```

```
FT      ACT_SITE      179      179      BY SIMILARITY.
SQ      SEQUENCE      317 AA; 34569 MW; CDDE4813AF6CB226 CRC64;

Query Match      39.8%; Score 647.5; DB 1; Length 317;
Best Local Similarity 44.2%; Pred. No. 1.2e-40;
Matches 137; Conservative 59; Mismatches 105; Indels 9; Gaps 5;

QY      18 KVVVGVGSGSATAYTLLLSGIVSEIVLIDVKNKDAEGESMDLNHA---APSNTRSRAG 74
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      8 KVVLLIGNGAVGSSYAFSLVNQSIVDLVIIDLTEKVRGDMDLKHATPSPPTTVRVKAG 67

QY      75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDITILLIATNPVD 134
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      68 EYSDCHDADLVVICAGAAQKPGETRLDLVSKNLKIFKFSIVGEVNMASKFDGIFLVATNPVD 127

QY      135 VLTYSYKASGFPPLSRVIGSGTGLDTRPFKYILGEHFKISSDSIDACVIGEHGD-GVPVW 193
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      128 ILAYATWKFSGLPKERVIGSGTILDSARFRLLSAEAFDVAPRSVDAQIIGEHGDTPLPVW 187

QY      194 SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIRKGYTSYSGIAAGLLRIVK 253
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      188 SHANIAGQPLKTLLEQRP--GKAQIEQIFVQTRDAAYDIIQAKGATYYGVAMGLARITE 245

QY      254 AILEDTGSTLTVSTV--GDYFGVEQIAISVPTKLNKSGAHQVAELSDEKEIELMEKSAS 311
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      246 AIFRNEDAVLTVSALLEGEY-DEEDVYIGVPAVINRNGIRNVVEIPLNDEEQSKFAHSAK 304

QY      312 QIKSVIEHLE 321
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db      305 TLKDIMAFAE 314

RESULT 13
LDH1_LACCA      STANDARD;      PRT;      325 AA.
ID      LDH1_LACCA
AC      P00343;
DT      21-JUL-1986 (Rel. 01, Created)
DT      01-DEC-1992 (Rel. 24, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      L-lactate dehydrogenase (EC 1.1.1.27) (L-LDH).
GN      LDH.
OS      Lactobacillus casei.
OC      Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC      Lactobacillus.
OX      NCBI_TaxID=1582;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 393 / DSM 20011 / NCDO 161;
RX      MEDLINE=92117575; PubMed=1768113;
RA      Kim S.F., Baek S.J., Pack M.Y.;
RT      "Cloning and nucleotide sequence of the Lactobacillus casei lactate
RT      dehydrogenase gene.";
RL      Appl. Environ. Microbiol. 57:2413-2417(1991).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 393 / DSM 20011 / NCDO 161;
RA      Taguchi H., Ohta T.;
RL      Submitted (NOV-1991) to the PIR data bank.
RN      [3]
RP      SEQUENCE.
RX      MEDLINE=83287369; PubMed=6411465;
RA      Hensel R., Mayr U., Yang C.;
RT      "The complete primary structure of the allosteric L-lactate
RT      dehydrogenase from Lactobacillus casei.";
RL      Eur. J. Biochem. 134:503-511(1983).
RN      [4]
RP      X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RA      Buehner M., Hecht H.J.;
RT      "Structure determination of the allosteric L-lactate dehydrogenase
RT      from Lactobacillus-casei at 3A resolution.";
RL      Acta Crystallogr. A 40:32-34(1984).
CC      -!- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) = pyruvate + NADH.
CC      -!- PATHWAY: Anaerobic glycolysis; final step.
```







Matches 136;		Conservative 59;	Mismatches 107;	Indels 8;	Gaps 5;
QY	18	KVVVVGVGVSATAYTLLSGIVSEIVLIDVNKDKAEGESMDLNHA---	APSNTRSRAG	74	
Db	7	KVVVIGTGFVGTSAIYSMINQGLVNELVLIDVNQDKAEGEALDLLDGVSWGQENVIVRAG	66		
QY	75	DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD	134		
Db	67	DYKCKNADIVVVTAGVNQKPGQSRLLDLVNTNAKIMRSIVTQVMDSGFDGIFVIASNPVD	126		
QY	135	VLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG-VPVW	193		
Db	127	ILTYVAVETSGLDQSRIVGTGTTLDTTRFRKELATKLEIDPRSVHGYIIGEHDSEVAVW	186		
QY	194	SLTNIDGMKLRDYCEKANHFQDNAFHRIFEQTRDAAYDIIKRKGYTSYGIAAGLLRIVK	253		
Db	187	SHTTVGGKPILEFIVNKKIGVED-LSNLSNKVNKNAAYEIIDKKQATYYGIGMSTARIVK	245		
QY	254	AILEDGTSTLTVSTV--GDYFGVEQIAISVPTKLNKSGAHQVAELSLDEKEIELMEKSAS	311		
Db	246	AILNNEQAILPVSAYLRGEX-GQEGVFTGVPISIVNQNGVREIIELNIDAYEKKQFEKSVS	304		
QY	312	QIKSVIEHLE	321		
Db	305	QLKEVIESIK	314		

Search completed: July 25, 2003, 11:44:13  
Job time : 24 secs









Db 183 FAVWSHANVAGVKLEQWFQE-NDYLNNEAEIVELFESVRDAAYSIIIAKKGATFYGVAVALA 241

QY 250 RIVKAILEDTGSTLTSTVSTVGD-YFGVEQIAISVPTKLNKSGAHQVAELSDEKEIELMEK 308

Db 242 RITKAILDDEHAVLPVSVFQDQGYGVSDCYLGQPAVVGAEGVNPPIHIPLNEAEMQKMEA 301

QY 309 SASQIKSVIE 318

Db 302 SGAQLKAIID 311

RESULT 7

P94886 PRELIMINARY; PRT; 325 AA.

AC P94886;

DT 01-MAY-1997 (TReMBLrel. 03, Created)

DT 01-MAY-1997 (TReMBLrel. 03, last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE L-lactate dehydrogenase (EC 1.1.1.27).

GN LDH.

OS Lactococcus lactis.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

OX NCBI\_TaxID=1358;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CM1-3;

RX MEDLINE=97176401; PubMed=9023947;

RA Urbach E., Daniels B., Salama M.S., Sandine W.E., Giovannoni S.J.;

RT "The ldh phylogeny for environmental isolates of Lactococcus lactis is consistent with rRNA genotypes but not with phenotypes.";

RL Appl. Environ. Microbiol. 63:694-702(1997).

CC -!- SIMILARITY: BELONGS TO THE LDH FAMILY.

DR EMBL; U78635; AAB51679.1; -.

DR HSSP; P00343; 1LLC.

DR InterPro; IPR001236; ldh.

DR InterPro; IPR001557; L\_LDH.

DR InterPro; IPR00205; NAD binding.

DR InterPro; IPR000594; ThiF\_domain.

DR Pfam; PF00056; ldh; 1.

DR Pfam; PF02866; ldh C; 1.

DR PRINTS; PR00086; LLDHDRGNASE.

DR PROSITE; PS00064; L\_LDH; 1.

KW NAD; Oxidoreductase.

SQ SEQUENCE 325 AA; 35038 MW; 1D1C9DED23EB12CB CRC64;

Query Match 36.1%; Score 586.5; DB 2; Length 325;

Best Local Similarity 40.0%; Pred. No. 3.3e-36;

Matches 124; Conservative 63; Mismatches 118; Indels 5; Gaps 4;

QY 13 DAKTVKVVVGVSGVSATAYTLTLLSGIVSEIVLIDVNMKKAEGESMDLNHAA--PSNTR 70

Db 3 DKQRKKVILVGDGAVGSSYAFALVNVQIAQELGIVDLFKEKTQGDADLSHALAFTSPKK 62

QY 71 SRAGDYPDCAGAAIVIVTCGINQKNGQTRMDLAARNANIMLEIIPNVAKYAPDPTILLIAT 130

Db 63 IYSADYSADSDADLVLTSGAPQKPGETRLDLVEKNLRITKDVVTKIVASGFKGIFLVAA 122

QY 131 NPVDVLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG- 189

Db 123 NPVDLTATYATWKESGFPKRVVSGTSLDTARFRQALAEKVDVDARSIHAYINGEHDSE 182

QY 190 VPVWSLTNIDGMKLRDCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYTSYGIAAGLL 249

Db 183 FAVWSHANVAGVKLEQWFQE-NDYLNNEAEIVELFESVRDAAYSIIIAKKGATFYGVAVALA 241

QY 250 RIVKAILEDTGSTLTSTVSTVGD-YFGVEQIAISVPTKLNKSGAHQVAELSDEKEIELMEK 308

Db 242 RITKAILDDEHAVLPVSVFQDQGYGVSDCYLGQPAVVGAEGVNPPIHIPLNEAEMQKMEA 301

QY 309 SASQIKSVIE 318

Db 302 SGAQLKAIID 311

RESULT 8

O81272 PRELIMINARY; PRT; 346 AA.

AC O81272;

DT 01-NOV-1998 (TReMBLrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Lactate dehydrogenase-2 (EC 1.1.1.27) (L-lactate dehydrogenase) (LDH).

GN LDH-2.

OS Solanum tuberosum (Potato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.

OX NCBI\_TaxID=4113;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Kennebec;

RA Dunford R., Ratcliffe R.G., Kruger N.J.;

RT "Lactate dehydrogenase (LDH) is regulated systemically at the level of gene expression, protein content and enzyme activity on flooding of potato plants.";

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

CC -!- CATALYTIC ACTIVITY: (S)-LACTATE + NAD(+) = PYRUVATE + NADH.

CC -!- PATHWAY: ANAEROBIC GLYCOLYSIS; FINAL STEP.

CC -!- SIMILARITY: BELONGS TO THE LDH FAMILY.

DR EMBL; AF067859; AAC21564.1; -.

DR HSSP; P00339; 9LDT.

DR InterPro; IPR001236; ldh.

DR InterPro; IPR001557; L\_LDH.

DR Pfam; PF00056; ldh; 1.

DR Pfam; PF02866; ldh C; 1.

DR PRINTS; PR00086; LLDHDRGNASE.

DR PROSITE; PS00064; L\_LDH; 1.

KW Glycolysis; NAD; Oxidoreductase.

SQ SEQUENCE 346 AA; 37628 MW; 02ED882ABEAA2515 CRC64;

Query Match 35.9%; Score 583; DB 10; Length 346;

Best Local Similarity 37.8%; Pred. No. 6.7e-36;

Matches 118; Conservative 76; Mismatches 110; Indels 8; Gaps 5;

QY 18 KVVVGVSGVSATAYTLTLLSGIVSEIVLIDVNMKKAEGESMDLNHAAPSNTRSR--AG 74

Db 36 KISVIGVGNVGMIAQITLTQDLVDELALVDVNSDKLGEMLDLOHAAAFPRTKIVASV 95

QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAARNANIMLEIIPNVAKYAPDPTILLIATNPVD 134

Db 96 DYTVTAGSDLCIVTAGARQNPGESRLNLLQRNLAMYKSIVPELVKYSPECILLIVSNPVD 155

QY 135 VLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHDG-GVPVW 193

Db 156 VLTYYVANK-SGFPVNRVIGSGTSLDSSRFRFLIADHLDVNAQDVQAYIVGEHGDSSVALW 214

QY 194 SLTNIDGMKLRDCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYTSYGIAAGLLRIVK 253

Db 215 SSISVGGIPVLSFLERQQIAFEKDTLEKHKQVVQVSAYEVINLKGYTSAIGYSVANLAF 274

QY 254 AILEDGTGTLTVS-TVGDYFGVE--QIAISVPTKLNKSGAHQVAELSDEKEIELMEKSA 310

Db 275 SIIRDQRRIHPVSILVKGFYIDGGDVFLSLPAQLGRSGVLGVTNVHLTDEEIQQLRNSA 334

QY 311 SQIKSVIEHLEI 322

Db 335 ETILEVQNQLGI 346

RESULT 9

Q96570 PRELIMINARY; PRT; 347 AA.

AC Q96570;

DT 01-FEB-1997 (TReMBLrel. 02, Created)

DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE L-lactate dehydrogenase (EC 1.1.1.27) (LDH).  
GN LDH.  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. UC82b;  
RX MEDLINE=98087999; PubMed=9426593;  
RA Germain V., Raymond P., Ricard B.;  
RT "Differential expression of two tomato lactate dehydrogenase genes in  
response to oxygen deficit.";  
RL Plant Mol. Biol. 35:711-721(1997).  
CC -|- CATALYTIC ACTIVITY: (S)-LACTATE + NAD(+) = PYRUVATE + NADH.  
CC -|- PATHWAY: ANAEROBIC GLYCOLYSIS; FINAL STEP.  
CC -|- SIMILARITY: BELONGS TO THE LDH FAMILY.  
DR EMBL; Y08888; CAA70101.1; -.  
DR HSSP; P00336; 5LDH.  
DR InterPro; IPR001236; ldh.  
DR InterPro; IPR001557; L\_LDH.  
DR Pfam; PF00056; ldh; 1.  
DR Pfam; PF02866; ldh C; 1.  
DR PRINTS; PR00086; LLDHDRGNASE.  
DR PROSITE; PS00064; L\_LDH; 1.  
KW Glycolysis; NAD; Oxidoreductase.  
SQ SEQUENCE 347 AA; 37659 MW; EF429D26A140738E CRC64;  
  
Query Match 35.8%; Score 582.5; DB 10; Length 347;  
Best Local Similarity 36.5%; Pred. No. 7.3e-36;  
Matches 114; Conservative 78; Mismatches 113; Indels 7; Gaps 4;  
  
QY 18 KVVVGVGSGVGFATYLLSGIVSEIVLIDVNVKDKAEGESMDLNHAAPSNTSR---AG 74  
DB 36 KISVIGVGNVGAIAQTILTQDLVDELALVDANSDKLRGEMLDLQHAAPFLPRTKIVASA 95  
  
QY 75 DYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNPVD 134  
DB 96 DYTVTAGSDLCIVTAGARQNPGESRLNLLQRLNLAAMYKSIVPELVKYSPECILLIVSNPVD 155  
  
QY 135 VLTYSYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHGD-GVPVW 193  
DB 156 LTTYVAVKLSGFVNRVIGSGTNLDSRRFRFLIADHLDVNAQDVQVIVGEHGDSSVALW 215  
  
QY 194 SLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKYTSYGIAGLLRIVK 253  
DB 216 SSISVGGIPVLSFLENQQAIEKDTLEKIHKKQVQVQSAYEVINLKGYTSWAIGYSVANLAF 275  
  
QY 254 AILEDGTSTLTSTVG-DYFGVE--QIAISVPTKLKNSGAHQVAELSLDEKEIEMKSA 310  
DB 276 SIIRDQRRIHPVSIKAGFYGIDGVDVFLSLPAQLGRNGVLGVTNVHLTDEEIQLRNSA 335  
  
QY 311 SQIKSVIEHLEI 322  
DB 336 ETILEVQNQLGI 347  
  
RESULT 10  
Q9W7M7 PRELIMINARY; PRT; 332 AA.  
AC Q9W7M7;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Lactate dehydrogenase A1 (EC 1.1.1.27) (L-lactate dehydrogenase)  
DE (LDH).  
GN LDHA1.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tsai S.C.M., Li J.Y., Mannen H., Li S.S.L.;  
RT "Molecular evolution of vertebrate lactate dehydrogenase isozymes by  
gene duplication.";  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
CC -|- CATALYTIC ACTIVITY: (S)-LACTATE + NAD(+) = PYRUVATE + NADH.  
CC -|- PATHWAY: ANAEROBIC GLYCOLYSIS; FINAL STEP.  
CC -|- SIMILARITY: BELONGS TO THE LDH FAMILY.  
DR EMBL; AF070952; AAD40730.1; -.  
DR HSSP; P00339; 9LDT.  
DR InterPro; IPR001236; ldh.  
DR InterPro; IPR001557; L\_LDH.  
DR Pfam; PF00056; ldh; 1.  
DR Pfam; PF02866; ldh C; 1.  
DR PRINTS; PR00086; LLDHDRGNASE.  
DR PROSITE; PS00064; L\_LDH; 1.  
KW Glycolysis; NAD; Oxidoreductase.  
SQ SEQUENCE 332 AA; 36460 MW; 045B897AA648399F CRC64;  
  
Query Match 35.8%; Score 581; DB 13; Length 332;  
Best Local Similarity 40.1%; Pred. No. 8.9e-36;  
Matches 125; Conservative 64; Mismatches 113; Indels 10; Gaps 5;  
  
QY 18 KVVVGVGSGVSATAYTLLSGIVSEIVLIDVNVKDKAEGESMDLNHAA----PSNTRSR 72  
DB 22 KVTIVGVGAVGMACAISVLQKDLADELALVDVIEDKLKGEWMDLQHGSFLRTPNIVSGK 81  
  
QY 73 AGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKNANIMLEIIPNVAKYAPDTILLIATNP 132  
DB 82 --DYSITANSKLVVVTAGARQEQEGESRLNVQRNVNIFKFIIPNVKYSKCTLLIVSNP 139  
  
QY 133 VDVLTYISYKASGFPLSRVIGSGTVLDTARFKYILGEHFKISSDSIDACVIGEHGD-GVP 191  
DB 140 VDILTYVAVKISGFPPQNRVIGSGCNLDSARFRLMGQKFGIHTQSCHGWVIGEHGDSSVP 199  
  
QY 192 VWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKYTSYGIAGLLRI 251  
DB 200 VWSGVNVAGVSLKSLHPDIGSDADKENWKEVHKQVDSAYEVIKLGKGYTSWAIGLSVADL 259  
  
QY 252 VKAILEDGTSTLTST-VGDYFGV-EQIAISVPTKLKNSGAHQVAELSLDEKEIEMKES 309  
DB 260 SESILKNLRRVHPISITVMKMGVGNEDVFLSVCVLGNLGITDVVTMTLKADEEEQLRKS 319  
  
QY 310 ASQIKSVIEHLE 321  
DB 320 ADTLWAIQKELQ 331  
  
RESULT 11  
Q96569 PRELIMINARY; PRT; 350 AA.  
AC Q96569;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE L-lactate dehydrogenase (EC 1.1.1.27) (LDH).  
GN LDH OR LDH1.  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. UC82B;  
RA Germain V., Raymond P., Ricard B.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. VFNB; TISSUE=Fruit;





QY 74 GDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKANANIMLEIIPNVAKYAPDTILLIATNPV 133  
Db 69 ATYADCADADLVITAGAPQKPGETRLDLVGKNLAINKSIVTQVVEGFGNGIFLVAANPV 128  
QY 134 DVLTYISYKASGFPLSRVIGSGTVLDTAREKYILGEHFKISSDSIDACVIGEHGDG-VPV 192  
Db 129 DVLTYSTWKFSGFPKERVIGSGTSLDSARFRQALADKIGVDARSVHAYIMGEHGDSEFAV 188  
QY 193 WSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYSYGAAGLLRIV 252  
Db 189 WSHANVAGVQLEQWLOE-NRDIDEQGLVDLFIISVRDAAYSIIKKKGATYYGIAVALARIT 247  
QY 253 KAILEDTGSTLTSTV--GDYFGVEQIAISVPTKLNKSGAHQVAELSDEKEIEMKSA 310  
Db 248 KAILDDENAVLPLSVYQEGYGDVKDFIGQPAIVGAHGIVRPVNIPLNDAELQKMQASA 307  
QY 311 SQIKSVIE 318  
Db 308 EQLKDIID 315  
RESULT 14  
Q8DZY3  
ID Q8DZY3 PRELIMINARY; PRT; 329 AA.  
AC Q8DZY3;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE L-lactate dehydrogenase.  
GN LDH OR SAG0959.  
OS Streptococcus agalactiae (serotype V).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=216466;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2603 V/R / Serotype V;  
RX MEDLINE=22222988; PubMed=12200547;  
RA Tettelin H., Massignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,  
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,  
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,  
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,  
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,  
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,  
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,  
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,  
RA Fraser C.M.;  
RT "Complete genome sequence and comparative genomic analysis of an  
RT emerging human pathogen, serotype V Streptococcus agalactiae";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).  
DR EMBL; AE014235; AAM99843.1; --  
DR TIGR; SAG0959; --  
KW Complete proteome.  
SQ SEQUENCE 329 AA; 35418 MW; 999D4BD9D9160B9A CRC64;  
Query Match 35.4%; Score 576; DB 16; Length 329;  
Best Local Similarity 41.2%; Pred. No. 2.1e-35;  
Matches 127; Conservative 58; Mismatches 115; Indels 8; Gaps 5;  
QY 18 KVVVVGVSATAYTLLSGIVSEIVLIDVKN--DKAEGESMDLNHAA--PSNTRSRA 73  
Db 9 KVLVGDGAVGSSYAFALVQNGIAQELGIIIEIPALFDKAVGDAEDLSHALAFTSPKKIYA 68  
QY 74 GDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKANANIMLEIIPNVAKYAPDTILLIATNPV 133  
Db 69 ATYADCADADLVITAGAPQKPGETRLDLVGKNLAINKSIVTQVVEGFGNGIFLVAANPV 128  
QY 134 DVLTYISYKASGFPLSRVIGSGTVLDTAREKYILGEHFKISSDSIDACVIGEHGDG-VPV 192  
Db 129 DVLTYSTWKFSGFPKERVIGSGTSLDSARFRQALADKIGVDARSVHAYIMGEHGDSEFAV 188  
QY 193 WSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYSYGAAGLLRIV 252

Db 189 WSHANVAGVQLEQWLOE-NRDIDEQGLVDLFIISVRDAAYSIIKKKGATYYGIAVALARIT 247  
QY 253 KAILEDTGSTLTSTV--GDYFGVEQIAISVPTKLNKSGAHQVAELSDEKEIEMKSA 310  
Db 248 KAILDDENAVLPLSVYQEGYGDVKDFIGQPAIVGAHGIVRPVNIPLNDAELQKMQASA 307  
QY 311 SQIKSVIE 318  
Db 308 EQLKDIID 315  
RESULT 15  
Q9W5Z7  
ID Q9W5Z7 PRELIMINARY; PRT; 332 AA.  
AC Q9W5Z7;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE L-lactate dehydrogenase M chain (EC 1.1.1.27) (LDH).  
GN LDH-A OR LDHA2.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Muscle;  
RA Mannen H., Li S.S.-L.;  
RT "Molecular evidence for a clade of turtles.";  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Tsai S.C.M., Li J.Y., Mannen H., Li S.S.-L.;  
RT "Molecular evolution of vertebrate lactate dehydrogenase isozymes by  
RT gene duplication.";  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: (S)-LACTATE + NAD(+) = PYRUVATE + NADH.  
CC -1- PATHWAY: ANAEROBIC GLYCOLYSIS; FINAL STEP.  
CC -1- SIMILARITY: BELONGS TO THE LDH FAMILY.  
DR EMBL; AF072582; AAD41639.1; --  
DR EMBL; AF070953; AAD40731.1; --  
DR HSSP; P00339; 9LDT.  
DR InterPro; IPR001236; ldh.  
DR InterPro; IPR001557; L\_LDH.  
DR Pfam; PF00056; ldh; 1\_LDH.  
DR Pfam; PF02866; ldh C; 1.  
DR PRINTS; PR00086; LDHDRGNASE.  
DR PROSITE; PS00064; L\_LDH; 1.  
KW Glycolysis; NAD; Oxidoreductase.  
SQ SEQUENCE 332 AA; 36472 MW; C293E0B4B79441B3 CRC64;

Query Match 35.3%; Score 574; DB 13; Length 332;  
Best Local Similarity 39.4%; Pred. No. 3e-35;  
Matches 123; Conservative 65; Mismatches 114; Indels 10; Gaps 5;  
QY 18 KVVVVGVSATAYTLLSGIVSEIVLIDVKNKKAEGESMDLNHAA-----PSNTRS 72  
Db 22 KVTIVGVGAVGMACAISVLQKDLADLALVDVIEDKLKGMMDLQHGSLFLRTPKIVSGK 81  
QY 73 AGDYPDCAGAAIVIVTCGINQKNGQTRMDLAAKANANIMLEIIPNVAKYAPDTILLIATNP 132  
Db 82 --DYSVTANSKLVVVTAGARQQEGESRLNLRQNVNIFKFIIPNIVKYSFNCTLLIVSNP 139  
QY 133 VDLTYISYKASGFPLSRVIGSGTVLDTAREKYILGEHFKISSDSIDACVIGEHGD-GVP 191  
Db 140 VDILTYVAWKISGFPKRVIGSGNLDLSARFRLMGQKFGIHTQSGHWIGEHGDSSVP 199  
QY 192 VWSLTNIDGMKLRDYCEKANHIFDQNAFHRIFEQTRDAAYDIIKRKGYSYGAAGLLRI 251  
Db 200 VWSGVNVAGVSLKTLHPDIGSDADKRWKVEVHKQVVDSEYEVIKLKYTSWAIGLSVADL 259

